CODING SHAPE INSIDE THE SHAPE

by

Rıza Alp Güler

Graduate School of Engineering and Natural Sciences

Master of Science Thesis

Sabancı University

Spring 2013-2014

Coding Shape Inside The Shape

APPROVED BY

Assoc. Prof. Dr. Gözde ÜNAL .....
(Thesis Supervisor)

Prof. Dr. Ali Rana ATILGAN .....

Prof. Dr. Aytül Erçil .....

DATE OF APPROVAL: 13.08.2014

©Rıza Alp Güler 2014
All Rights Reserved

Acknowledgements

I am very grateful to the many people who have helped and inspired me during my masters study.

I owe many thanks to my thesis supervisor Prof. Gözde Ünal for believing in me. I have learned a lot from you. You have always given me the opportunity to be free and you have always made yourself accessible when I needed help. You have guided me throughout the thesis. Thanks to your sincerity and kindness, I always felt very comfortable throughout my study, even at strenuous times.

I am very thankful to my professors, whose lectures have greatly inspired me. Prof. Aytül Erçil, who brilliantly and joyfully reflects her experience onto her students , Prof. Müjdat Çetin, who showed me how clear complicated things can get once you do things precisely, Prof. Hakan Erdoğan, who showed me it is possible to teach very effectively by implementing in class. I am also grateful for all my friends at VPA Lab, their company will never be forgotten.

I would like to thank Prof. Sibel Tari. Her work was a great motivation for this thesis.

I thank my parents and brother for their unconditional love and support.

Thank you, Gizem. Less is more; more or less.

I would like to thank TÜBİTAK for providing financial support for my graduate education. I was supported by TÜBİTAK 1001 Grant No: 112E320 : "PrePostOp-DTI: New Mathematical Computing Techniques for Analysis of Pre-Operation vs Post-Operation Changes in Brainstem White Matter Tracts using Diffusion Tensor Imaging".

Coding Shape Inside The Shape

Rıza Alp Güler

EE, M.Sc. Thesis, 2014

Thesis Supervisor: Gözde ÜNAL

Keywords: Shape Analysis, Shape Representation, Shape Coding, Elliptic models for Distance Transforms, Scalable Fluctuating Distance Field, Screened Poisson Hyper-Field, Local Convexity Encoding Field, Shape Decomposition, Non-Rigid Shape Retrieval

Abstract

The shape of an object lies at the interface between vision and cognition, yet the field of statistical shape analysis is far from developing a general mathematical model to represent shapes that would allow computational descriptions to express some simple tasks that are carried out robustly and effortlessly by humans. In this thesis, novel perspectives on shape characterization are presented where the shape information is encoded inside the shape. The representation is free from the dimensions of the shape, hence the model is readily extendable to any shape embedding dimensions (i.e 2D, 3D, 4D). A very desirable property is that the representation possesses the possibility to fuse shape information with other types of information available inside the shape domain, an example would be reflectance information from an optical camera.

Three novel fields are proposed within the scope of the thesis, namely 'Scalable Fluctuating Distance Fields', 'Screened Poisson Hyperfields', 'Local Convexity Encoding Fields', which are smooth fields that are obtained by encoding desired shape information. 'Scalable Fluctuating Distance Fields', that encode parts explicitly, is presented as an interactive tool for tumor protrusion segmentation and as an underlying representation for tumor follow-up analysis. Secondly, 'Screened Poisson Hyper-Fields',

provide a rich characterization of the shape that encodes global, local, interior and boundary interactions. Low-dimensional embeddings of the hyper-fields are employed to address problems of shape partitioning, 2D shape classification and 3D non-rigid shape retrieval. Moreover, the embeddings are used to translate the shape matching problem into an image matching problem, utilizing existing arsenal of image matching tools that could not be utilized in shape matching before. Finally, the 'Local Convexity Encoding Fields' is formed by encoding information related to local symmetry and local convexity-concavity properties.

The representation performance of the shape fields is presented both qualitatively and quantitatively. The descriptors obtained using the regional encoding perspective outperform existing state-of-the-art shape retrieval methods over public benchmark databases, which is highly motivating for further study of regional-volumetric shape representations.

ŞEKİL İÇERİSİNE ŞEKİL KODLAMA

Rıza Alp Güler.

EE, Yüksek Lisans Tezi, 2014

Tez Danışmanı: Gözde ÜNAL

Anahtar Kelimeler: Şekil Analizi, Şekil Temsili, Şekil Tanıma, Şekil Bilgisi, Şekil Yakınlığı, Şekil Eşleme, Tümör Şekil Analizi, Sönümlü Poisson Hiper-Alanları, Ölçeklenebilir Dalgalı Mesafe Alanları, Bölgesel Konveksite İşleyen Alanlar

Özet

İnsan beyninde görmek ve algılamak arasında gerçekleşen henüz tam belirli olmayan bir süreçle tanımlandırılan 'şekil' için mevcut matematiksel modeller, halen insanların kolayca çözdüğü tanımlama problemlerinin çözülmesini sağlayacak temsiliyeti sağlayamamaktadır. Bu tez kapsamında şekil tanımlama ile ilgili yeni bir bakış sunulmaktadır. Bu yeni şekil tanımlaması (temsili), şekil ile ilgili hesaplanan bilgileri şeklin içerisine kodlanması ile oluşturulmaktadır. Elde edilen şekil temsili, şeklin boyutu ile alakalı değildir. Bu yönüyle önerilen model farklı boyutlardaki şekiller için geçerlidir (ör. 2B, 3B, 4B). Yaratılan temsilin bir başka önemli özelliği ise, şekil ile ilgili niteliklerin bölgesel olarak erişilebilir olması sebebiyle, şekil bilgisinin şekil üzerinde tanımlanmış olan başka türden bilgilerle birlikte kullanılmasına elverişli olmasıdır. Tez kapsamında üç farklı şekil temsil yöntemi önerilmektedir: Sönümlü Poisson Hiper-Alanları, Ölçeklenebilir Dalgalı Mesafe Alanları, Bölgesel Konveksite İşleyen Alanlar. Önerilen şekil nitelendirme yöntemleri görsel sonuçların yanında çeşitli uygulamalarda sayısal sonuçlar ile sunulmaktadır. Sunulan uygulamalardan bazıları: şekil parçalama, şekil sınıflandırma, şekil yakınlığı belirleme, şekil eşleştirmesi ve tümör şekli çakıştırmasıdır. Sayısal sonuçlar önerilen bölgesel temsil yöntemlerinin bazı problemlerde bütün modern metotlardan daha gürbüz ve başarılı çalıştığını göstermektedir.

Table of Contents

Acknowledgements **iv**

Abstract **v**

Özet **vii**

1 Introduction **1**
1.1 On Shape Analysis . . . 1
1.1.1 Links to Human Perception of Visual Form . . . 2
1.1.2 On Shape Representation . . . 3
1.2 Contributions and Thesis Outline . . . 6
1.2.1 *Scalable Fluctuating Distance Fields* . . . 6
1.2.2 *Screened Poisson Hyper-Fields* . . . 7
1.2.3 *Local Convexity Encoding Fields* . . . 9

2 A Scalable Fluctuating Distance Field **11**
2.1 A Part-Based Representation for Tumor Shapes . . . 11
2.1.1 Related Work . . . 12
2.1.2 Our Contribution . . . 13
2.2 Scalable Fluctuating Distance Field . . . 14
2.2.1 Energy Terms . . . 15
2.2.2 A Sign Constraint to Control Fluctuation Scale . . . 16
2.2.3 A Space of Fluctuation Scales . . . 18
2.2.4 Interactive Tumor Protrusion Segmentation . . . 21
2.3 Tumor Follow-Up Registration Using ω fields . . . 22
2.3.1 Registration Results and Discussion . . . 24
2.4 Conclusions . . . 25

3 Screened Poisson Hyper-Fields **29**
3.1 Introduction . . . 29
3.1.1 Related Works . . . 29
3.1.2 Our Contribution . . . 33
3.2 A new hyper-field . . . 35
3.2.1 A two-dimensional Scale Space . . . 35
3.2.2 Varying ρ^2: Sweeping Internal Smoothing Characteristics . . . 36

3.2.3 Fixing ρ^2: Decomposition of Boundary Sources . . . 37
3.2.4 Putting it altogether: The New Hyper-field . . . 40
3.3 Screened Poisson: Properties . . . 41
3.3.1 Screened Poisson as a conditioned random walk . . . 41
3.3.2 Relation to Geodesic Distances . . . 43
3.3.3 Relation to Spectral Methods . . . 44
3.4 Extracting information from Hyper-fields . . . 46
3.4.1 Unveiling parts from the hyper field via sparse coding . . . 46
3.4.2 Producing consistent mappings for shape correspondence: SPEM . 49
3.5 Results and Discussions . . . 53
3.5.1 Computational Aspects . . . 53
3.5.2 Boundary Decomposition Based on Regional Information . . . 54
3.5.3 Orthogonal Projections Based On ρ^2 Sweep: SPEM . . . 55
3.5.4 A Moment Based Evaluation of Consistency and Correspondence . 60
3.5.5 Non-Rigid Shape Retrieval Using Screened Poisson Encoding Maps 62
3.6 Conclusion . . . 67

4 Shape Matching using Image Descriptors 69
4.1 Introduction . . . 69
4.1.1 Related Works . . . 69
4.1.2 Our Contribution . . . 70
4.2 2D Shape Matching and Retrieval Using Consistent Projections . . . 71
4.2.1 Shape Retrieval Approach . . . 75
4.2.2 Shape Retrieval Experiments . . . 76
4.2.3 Discussion . . . 80
4.3 Conclusions . . . 81

5 Local Convexity Encoding Fields 83
5.1 A new gaze into the hyper-field . . . 83
5.2 Local convexity encoding fields . . . 85
5.3 Discussion and Conclusion . . . 89

6 Conclusions and Future Directions 92

Bibliography 93

List of Figures

2.1 *Left* The normalized field $\omega(x = \tilde{x}, \eta)$, where $\tilde{x}$ is shown by horizontal(top) and vertical(bottom) red lines. Image obtained by sweeping η from 1 to -1. *Right* Surface plot for $\omega(x, y, \eta) = 0$ 19

2.2 *From left to right:* Input shape, ω for $\eta > 0$, ω for $\eta = 0$, ω for $\eta < 0$ . . . 19

2.3 Ω^- domain and watershed segmentation results for: *left* $\eta > 0$, *right*: $\eta = 0$ 20

2.4 The original ω field(left)[1], where the Lagrange multiplier γ is chosen as zero in Eq.2.12 and five ω fields (right) calculated using increasing values for η, where $\eta_1 < \eta_2 < \eta_3 = 0$ and $\eta_5 > \eta_4 > \eta_3 = 0$. Upper row for both shapes is a contour plot of normalized ω and bottom rows depict ω for solely Ω^-. 21

2.5 *Left:* An axial slice of contrast enhanced T1 MRI of a patient with a tumor. *Middle:* ω field isocontours for the corresponding tumor slice. *Right:* ω field visualized. 22

2.6 *From left to right:* Tumor volume. Positive and negative parts of the proposed field. Positive part of the field. Negative part of the field. Segmented protrusions of the tumor enveloped in the negative part of the field. Segmented protrusions visualized with the positive part of the field. 22

2.7 Visualizations of positive(opaque) and negative(transparent) parts of the tumor field paired with corresponding segmentation results. The fluctuating distance field for each pair were generated using the corresponding η value. 23

2.8 (a,b,c): Visualizations of deformation field vectors and volume change pairs for registration of each synthetic shape couples, generated using Left:Distance Transforms Middle: Normalized Distance Transforms. Right: Scalable Fluctuating Distance Fields. 27

2.9 For both parts of the figure: *Left:* Displacement field vectors from gray initial tumor to blue followup tumor. *Middle:* The displacement vectors to a specific segmented protrusion. *Right:* Local volume change maps in initial tumor domain for selected axial slices of the tumor shapes, the black contours denote the followup tumor. The maps on the left and right are generated from the deformation fields calculated using normalized $\mathcal{D}$ and ω fields respectively. 28

3.1 v fields for different values of ρ^2. 35

3.2 Field value versus ρ at five selected nodes of distinct characters. v function is coding characteristics that extend beyond usual distances. A dense linear sampling is used between $\rho = 2$ and $\rho = 30$. 37

3.3 Behavior of v^ρ in ρ dimension for sampled points on the domain that are equidistant to the boundary. 38

3.4 Solutions of the screened Poisson equation for a 1D experiment using three different boundary conditions (columns) and three different ρ values (rows). 39

3.5 Restricting the boundary inhomogeneity to a single point $\boldsymbol{p_i}$ on the little finger. a) Iso-contours (bottom) and values of v-field using $log(v^{\boldsymbol{p_i}})$ visualised as a point cloud; b) Normalized gradient, $\frac{\nabla v \boldsymbol{p_i}}{|\nabla v \boldsymbol{p_i}|}$ for the 'thumb'; c) Streamlines obtained by tracking along the normalized gradient directions. 40

3.6 Separating sources of variability in the shape hyper-field. 41

3.7 Non-negative measurements: $y_j(i)$, where the same $\boldsymbol{p_i}$ in Fig. 3.5 is used. *Left:* Normalization by *median*. *Right:* Normalization by *mean*. 47

3.8 a) NNSC components obtained using a large λ and $k = 5$; b) NNSC components obtained using a low λ and $k = 12$. 48

3.9 For the hand shape: Left: calculated Φ_n colored according to corresponding scale ratios; Right: PSNR values for projections obtained using Φ_n across different shape scales show a slow monotonically changing behaviour, which provides a desired robustness to scale changes, color coded as shown on the bottom right. 51

3.10 *Top:*$\{n/4\}$regular star polygons, for $n = 9, .., 20$ *Left:* First six eigenvectors $\Phi_1, ..\Phi_6$ for the shapes colored accordingly. *Right:* The first six eigenvectors $\Phi_1, ..\Phi_6$ for the shapes after re-scaling with respect to the maximum value of the shortest distance to the boundary. 52

3.11 Decomposition of the human figure and associated regions. $k = 8$. 54

3.12 Non-negative sparse decomposition over shape hyper-fields of three highly different cat poses partition shape boundary into: the head, the cat frontal body, the back body, its tail, and its legs in a consistent manner. 55

3.13 Left-Right: First five projections(SPEM): $\mathcal{P}_{1,...,5}$ for 6 different poses of a cat shape, depicted in each row. Each column corresponds to a different projection mode. Hotter colors indicate positive and high values while colder colors indicate negative and low values. Consistency of projections across deformations of the cat shape is observed. 56

3.14 First six projections(SPEM): $\mathcal{P}_{1,...,6}$ (on each row) for five different instances (on each column) of a human and a hand silhouette. Human figure displays articulated motion and local deformations. Hand figure displays different noise conditions: occluding a finger; shortening of fingers; protruding two new parts from the hand. Hotter colors indicate positive and high values while colder colors indicate negative and low values. Robustness of projections against occlusion, local deformation, and noise is observed. 57

3.15 Top-Down: Second to sixth projections(SPEM): $\mathcal{P}_{2,...,6}$ for three different poses of a 3D horse. Consistency of each projection across a row for different poses can be observed. 59

3.16 Each row contains the negative-positive nodal domain clusters for corresponding to first 5 projections of 7 cat shapes. 61

3.17 SVM Classification accuracies using moment features of: binary shape moments (blue); individual thresholded projections (red); and cumulatively adding thresholded projections (black). Notice that the success rate jumps from %30 (blue) to %80 (blacks) when our approach is used. 62
3.18 Joint histograms inside SPEMs: $\mathcal{P}_4$ vs $\mathcal{P}_3$ for corresponding shapes on the right. The histogram intensities are displayed using a logarithmic scale. The articulations have almost no effect on the joint histograms and there is large variation in histograms of shapes with different volumetric structures. 64
3.19 Precision - Recall Performances in Shrec'11 Non-Rigid Database 67

4.1 (a) 2NN SIFT Matching results for $\mathcal{P}_i, i = 1, 2, .., 6$ for two cat shapes. (b,c) Refined matches after geometric verification with a (b)*relaxed* (c) *strict* planar homography assumption. For solely visualisation purposes, the matches are clustered using a k-means algorithm based on spatial distances. 73
4.2 Matching results for four pair of shapes. *First Row:* Input and target shapes. *Second Row:* Inlier matches within planar homography. *Third Row:* Input shape projected onto the target using the estimated homography. . 75
4.3 Proposed Shape Matching Scheme 75
4.4 Distributions of match contributions versus inlier ratios from each projection. Left: Matches between shapes of different categories. Right: Matches between shapes of same category. 76
4.5 Precision Recall graphs for retrieval experiments on three datasets. 80

5.1 Projections using first two principal components of $log(v(x_2, y_2, \rho))$ on various shapes of closed contours. 84
5.2 $V_{Ref}(\mathcal{D}(\vec{x}))$ vs time as reference radius increases. 85
5.3 $\Psi(\vec{x}, t)$ and partitioning by the zero level curve are presented for two cat shapes in five time instances. Time instances are matched by the minimum vale of $v_\rho(\vec{x}, t)$. The values used are 0.1 for t_1, 0.25 for t_2, 0.5 for t_3, 0.75 for t_4 and 0.9 for t_5. 86
5.4 $A_{final}(\vec{x})$ fields and extracted parts for various shapes. 87

5.5 Projections using the principal components of the hyper-field $A(\vec{x}, t)$ in the time domain for two shapes. 89

5.6 $A(\vec{x}, t)$ at two different t values, for the shape presented in Attneave's work[2]. Subjects attempted to approximate the closed figure shown above with a pattern of 10 dots. Radiating bars indicate the relative frequency with which various portions of the outline were represented by dots chosen. Note how the A hyperfield simulates the human subjects in marking convex and concave points on the boundary of the shape. 90

List of Tables

3.1 SHREC'11 Retrieval on Non-Rigid 3D Watertight Meshes Database Results 66

4.1 Retrieval Performance for 490 shape database for 15 closest shapes. 78
4.2 Retrieval Performance for 180-shape dataset 79
4.3 Retrieval Performance for 1000-shape dataset 80
4.4 Retrieval Performance using context methods for the 1000-shape database. 80

1

Introduction

Perception of visual information occupies about 70% of our cortex activity. Apparently, about the same percentage of visual content occupies communication networks over the world[3]. Geometric information regarding the *shape* of objects constitutes a large portion of the visual information and has been studied for decades by scientists from many different fields.

1.1 On Shape Analysis

Shape analysis is currently playing a pivotal role in many applications from a variety of fields. It has become one of the major topics in the field of Computer Vision. Characterization of complex objects using their global shapes is fast becoming a major tool in computer vision and image understanding. [4], eg. applications like classification of objects or content-based object retrieval. Introducing shape information to problems in computer vision is very desirable, considering other sources of information like reflectance, lighting, texture can get quite uninformative. With improvements in sensor technologies, it is possible to obtain shape information from RGB-D cameras, which helps discarding the segmentation problem and provides valuable shape information. Another field of study that is highly connected to shape analysis is medical image analysis. Some existing approaches and current problems are presented in Gao's work[5]. Various neurodegenerative and neurodevelopmental brain disorders are successfully linked to brain

morphometry[6, 7, 8, 9, 10, 11, 12, 13, 14]. Shape analysis has also been of great interest to computer graphics community, where shapes are analyzed as 3D boundary meshes in applications involving shape segmentation/partitioning[15], shape retrieval[16], shape correspondance[17], and so on. Shape analysis techniques has been utilized in many other fields, some of which are paleanthology[18], archeology[19] and biology[20]. For a more detailed, in-depth summary of existing problems concerning shape analysis and its applications, the readers are referred to books by Small[21], Dryden and Mardia[22], Krim and Yezzi[23], Kendall, Barden and Carne[24].

1.1.1 Links to Human Perception of Visual Form

Research on shape analysis has been greatly influenced by the human perception of visual form. Findings on human visual system by disciplines of psychology, cognitive science, art and more recently neuroscience[25] have motivated some of the seminal works in the field. According to Pylyshyn[26](and many others), analysis of shape lies in interface between vision and cognition as a part of the *early vision* system. Outlining existing relations or the parallelism between human cognition and shape analysis literature is certainly out of the scope the thesis, yet due to crucial relevance, the seminal work is introduced following the differentiation: classical - modern theories of visual perception by Loncaric[27].

The revolutionary Gestalt school of psychology[28, 29, 30] provides principles (laws) on properties of visual forms. The central principle of Gestalt psychology is that the mind forms a global whole with self-organizing tendencies. Even though the parts change the whole can remain unaltered. Koffka portrays this by saying "the whole is other than the sum of its parts". Other classical theories of visual form are Hebb's theory[31], where form is not perceived as a whole but consists of parts, and theory of Gibson[32], where monocular cues(stimulus) like texture, saturation of colors, shading, parallel lines are used in perceiving real three-dimensional objects. The latter is in contrast to the Gestalt theory, where the dynamism of real world objects is analyzed as an ambiguity of the interpretation of images projected into a two-dimensional space. These classical theories of visual form are non-computational. This aspect poses a disadvantage for practical

engineering applications[27].

As for modern theories of the human visual perception system, in Marr's work[33, 34] the focus of research is shifted from applications to topics corresponding to modules of the human visual system. The computationally supported work opened new directions to the field of shape analysis[27], specifically to the concept: *shape from x*, which deal with reconstructing shape from cues including shading[35, 36, 37], stereo[38], texture[32, 39], contour[40], focus[41], etc. Lowe[42] with a similar motivation, introduced methods for recognition of three-dimensional objects from unknown viewpoints solely using a two-dimensional image. A *dynamical shape model* was proposed by Koenderink et al.[43, 44], where on several scales of resolution was considered for the modeling of visual perception. Such a *hierarchical* representation for shapes was also considered in art before and the idea is in the core of many recent successful computational approaches of shape representation. Attneave performed psychological experiments[2] to demonstrate that visual data is highly redundant and portrays that points of high curvature on shape boundaries are informative and perceptually relevant.

1.1.2 On Shape Representation

Representation of the shape is in the heart of any shape analysis approach. A large number of shape representation techniques are proposed to address different problems. A shape representation is modeled with a preconception of the application and invariance properties desired. The fact that there is no verbal definition of "Shape" that applies for every shape analysis scenario, is highly coherent with the large quantity of existing shape representations.

Kendall's definition[45] of shape: *"all the geometrical information that remains when location, scale and rotational effects are filtered out from an object."* is well acknowledged, yet there is no default invariance group to address all problems. For instance, rotation invariance is not desired in a optical character recognition task, since letters p and d would be identical according to the representation. The representation is expected to be invariant to bending or articulated motion for applications like retrieval of perceptually similar shapes. On the other hand, bending invariance and scale invariance is not desired in

analysis of anatomical organs, because bending and global scale change could be symptoms of interest.

Shape analysis of 2D silhouettes as perspective projections of real world 3D objects is of great interest for the computer vision community. Representation in this case becomes crucial because one dimension of geometric information is lost. Real world objects undergoing slight pose changes might cause severe topological changes in the projected planar shape, which only partially represents the actual object. Additionally effects like occlusion, distortion and noise further complicate the problem. The problem of noise also occurs in the field of medical imaging, where some of the medical data acquisition methodologies provide insufficient resolution. A shape representation technique should be robust against these undesired effects and should be equipped with desired invariance properties. The representation is commonly used to declare a shape similarity measure, some examples are [46, 47, 48, 49, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59], which are utilized to retrieve perceptually similar shapes from a database to a given query shape or to evaluate the quality of various medical imaging tasks such as image registration and segmentation as in [60].

An earlier classification of shape representation techniques was made by Pavlidis, [61]. Zhang et al.[62] also follows classification of Pavlidis as *boundary* and *internal(region based)* shape representation techniques. A more recent classification is presented as a survey by Yang et.al.[63]. Boundary based approaches are contour-based shape representation techniques that exploit shape boundary information. The information is generally encoded in the form of a string, tree or graph, so that a similarity measure can be extracted using string or graph matching techniques. In the earlier classification [61, 62],the regional shape representation techniques are mainly centered around the medial axis transform(MAT) by Blum[64, 65]. The revolutionary idea of Blum is representing the shape using the local symmetry axes, which encode regional characteristics of shapes. The non-computational idea of Blum was interpreted(computationally) by many different methodologies including morphological operations and voronoi cells, but perhaps the best computational analogy[66] to Blum's *prairie-fire* definition is the level-set methods, which will be described next. For a more detailed analysis of general shape representation techniques, especially on

boundary-based approaches the readers are referred to the surveys[61, 62, 63]. Some of the shape representation approaches that have relevance for the work will be introduced more extensive throughout the thesis .

In the 90's, it was observed that shapes can be embedded as zeros of a function defined over the shape domain, opening the way to an active research area in implicit shape representations ([67], [68]). In this area of research, the signed distance transform was popularized heavily by the level-set framework [69] and its fast implementation [70].

The distance function is created via solution of the Eikonal equation $|\nabla u(x)| = 1, x \in \Omega$ subject to boundary condition $u|_{\partial\Omega} = 0$. The governing equation forces the absolute value of the gradient to be constant. Equipped with a suitable boundary condition, the solution $u(x)$ is interpreted as the shortest time needed to travel from the boundary to the point x. Signed Distance Transform (SDT) is formed by setting positive and negative of the distances exterior and interior to the shape or vice versa, facilitating regional encoding of shape domain and its exterior by minimal distances to the shape boundary. The shape is then represented as the zero level set of the signed distance transform (SDT). This representation of the shape, *i.e.* via embedding the shape boundary as the level set of SDT, became quite instrumental in developing approximate schemes for segmentation functionals and introducing shape knowledge in segmentation problems[71].

Perhaps the most significant property of the level-set representation is that the information regarding shape is in a space where other sources of information can be reached, thus integrated to the problem. This has been utilized in segmentation applications where shape information is fused with other types of low-level information such as edge consistency [72, 73, 74], intensity homogeneity[75, 76], texture information[77, 78, 79, 80, 81] and motion information[82, 83]. For a more detailed analysis on how information regarding color, texture, motion and shape are integrated to the framework, the readers are referred to the survey[84]. Work of Tari et al.[85] provides an alternative regional encoding approach, similar to the reaction-diffusion process, used by Kimia et al. [66]. A field inside the shape is obtained whose iso-contours mimic curvature dependent evolution. The field is used for extracting skeletons from gray-scale images, which is a sound example on how shape information and other cues can be tied thanks to the regional distance-like

underlying representation.

1.2 Contributions and Thesis Outline

Within the scope of the thesis, three novel shape representation methods are introduced, which share a common property:

> Information regarding shape, extracted from internal distance relationships, is encoded **inside** the shape domain in a smooth manner.

Since the information is coded inside the shape, it is the possible to fuse shape information with other types of information available inside the shape domain. For instance, reflectance information regarding an object obtained from an optical camera or information regarding water content, functional activity or diffusion characteristics from various MRI acquisition methods could be fused with shape information. Unlike the SDT used in level-set methods, our fields are very informative, as we will show throughout the thesis, so such information could be useful in various tasks including characterization. Another acknowledged advantage of the presented shape representation, which is valid for all distance-based shape representations, is that the shape model is free from the dimensions of the shape: the volumetric description of the hyper-fields we propose is readily extendable to any shape embedding dimensions in $\mathbb{R}^n$ for $n = 2, 3, 4....$

Chapters in this thesis include an analysis of existing work that has relevance for the content of the chapter and conclusions regarding the content of the chapter. Next, the contents of the chapters are described. Finally, at the end of the thesis, conclusions in a general manner and remarks on future work will be stated.

1.2.1 *Scalable Fluctuating Distance Fields*

The first shape encoding field is designed with a motivation to represent tumor shapes. Tumor growth involves highly complicated processes and complex dynamics, which typically lead to deviation of tumor shape from a compact structure. Motivated from physical significance and clinical relevance in follow-up problems, we proposed a method to analyze

the protruded and peripheral regions of tumor shapes. The modified field is introduced in $\mathbb{C}.2$[1], along with analysis of relevant work.

In the earlier work of fluctuating distance fields [1], the shape field consists of positive and negative values whose zero crossing separates the central and the peripheral volumes of a silhouette. We add a non-linear constraint upon the original fluctuating field idea in order to introduce a "fluctuation scale", which indicates an assumption about peripherality. This provides the induction of an hierarchy hypothesis onto the field. By varying the fluctuation scale from low to high values, it is possible to observe the coarse to fine levels of hierarchy both in the field and its segmentations even by utilizing a very simple segmentation method. We discuss the scale-space arising from the aditional parameter. When the parameter is fixed, the field becomes robust for scale changes for analysis of correspondence, albeit the loss of the linearity of the original shape field model.

The proposed modification leads to an interactive framework for segmenting the protrusions and partitioning tumorous structures In order to quantify the tumor shape variations in a follow-up scenario, a shape registration based on a scalable fluctuating shape field is described. The representation performance of the scalable field for a fixed 'fluctuation scale' is demonstrated in comparison to the conventional distance transform approach for the registration problem. The scalable shape field becomes a potentially powerful underlying shape representation for shape registration procedures, due to an increased robustness to scale changes without losing the information it inherits particularly in terms of the parts of a shape.

1.2.2 *Screened Poisson Hyper-Fields*

The second regional representation presented in the thesis is the *shape hyper-fields*. This is a novel perspective on shape characterization using the screened Poisson equation, which was first used for disconnected skeleton extraction from shapes. We discuss that the effect of the screening parameter is a change of measure of the underlying metric space; also indicating a conditioned random walker biased by the choice of measure. A continuum

[1] *Scalable Fluctuating Distance Fields* is published at Springer Book Series: Research in Shape Modelling

of shape fields is created, by varying the screening parameter or equivalently the bias of the random walker. In addition to creating a regional encoding of the diffusion with a different bias, we further break down the influence of boundary interactions by considering a number of independent random walks, each emanating from a certain boundary point, and the superposition of which yields the screened Poisson field. Probing the screened Poisson equation from these two complementary perspectives leads to a high-dimensional hyper-field: a rich characterization of the shape that encodes global, local, interior and boundary interactions. We discuss two low-dimensional embedding schemes, one to unveil parts using non-negative sparse coding[86] and the other to produce consistent mappings, which we call Screened Poisson Encoding Maps (SPEM), for the purpose of shape matching and shape retrieval. Details regarding the Hyper-Fields are given in $\mathbb{C}$.3[2].

The potential of extracting various shape descriptors from the introduced shape hyper-field was demonstrated over both a 2D "1000-shape" database with a moment based approach. For 3D non-rigid shape retrieval, we use the benchmark dataset *SHREC'11* [87] . The SPEM performance was evaluated by using the VLAD method [88] for volumetric feature encoding. The SPEM consistently ranked the first or the second in all measures, and ranked the first when a hybrid combination with top surface-based methods was computed. The results of SPEM suggest that extracting volumetric information in a robust way can lead to enhancement in the non-rigid shape retrieval performance when compared to extracting information regarding only intrinsic surface properties, which is very motivating for further study of volumetric representations.. Moreover, as expected, combining volumetric information and surface information results in a significant boost in performance in 3D shape retrieval.

Contrasting to boundary-based approaches, our shape-interior based representation allows the landmarks to be obtained from the whole shape domain, which leads to robustness to artifacts that can occur in shape boundaries. In addition, correspondences are obtained by combination of *almost local* (shape field measurements) and *global* (eigenvectors) cues due to the characteristics of the projections. The fact that SPEM's are robust

[2] *Screened Poisson Hyper-Fields: A New Perspective In Shape Representation* is currently in second revision round for SIAM, Journal on Imaging Sciences

to scale changes and show consistency across various scenarios: different shape poses, deformations, occlusions and clutter, motivates their use in 2D shape matching. Yet using SPEM's for nodes as features or VLADS is not enough considering the alterations of regional shape characteristics due to projective transformation.

In the last few decades, significant advances in image matching are provided by rich local descriptors that are defined through physical measurements such as reflectance. As such measurements are not naturally available for silhouettes, existing arsenal of image matching tools cannot be utilized in shape matching. We advocate use of SPEM's to translate shape-matching problem into image-matching problem. We devise a shape similarity measure based on the SIFT[89] descriptors of the projections, which is later utilized in a matching scheme refined by RANSAC[90] to yield state-of-the-art retrieval results. Details regarding this method is given in $\mathbb{C}.4$[3].

Thanks to both the holistic and regional nature of the provided shape representation, a SIFT-based image matching framework, could be used in 2D shape matching, for the first time to our knowledge. Even the surprisingly simple idea employed as the shape similarity measure, which is the total number of retained correspondences across corresponding shape projections of the two shapes being compared, achieves very good performance in matching as demonstrated by the performance over three common shape datasets. The presented shape matching scheme performs favorably among some popular shape retrieval methods.

1.2.3 *Local Convexity Encoding Fields*

Finally, we propose a shape field that encodes convexity and concavity inside the shape domain. The motivation is noticing that most of the variance in the hyper-field is related to the distance of the nodes to the boundary. We estimate a reference field by forming a relationship between the distance of a node to boundary, i.e. the distance transform ($\mathcal{D}$) and the solution to the poisson equation. The proposed field is formed by aggregating the deviation of the field from the reference field in time. The reference field is is modeled as an answer to the following question for a node at point ($\vec{x}$):

[3]Manuscript *SIFT for Shapes* under preparation

Considering the distance of the node to boundary, $\mathcal{D}(\vec{x})$, how different would the intensity of node ($v(\vec{x}, \rho)$) be at time t if the source(shape boundary) was a perfect circle.

The new field directly separates external parts from the central region. The isocontours of the field reveals that rich shape information is encoded, when potential just arising from distance to source is discarded. The partitioning of the shapes is obtained using a very natural and perceptually coherent manner. A new hyper-field is formed from the evolution of this field in time, whose projections give a hint of the information encoded.We observe medial loci of the shapes and concave regions explicitly. The results are also consistent for the shapes of the same class. The field is introduced in $\mathbb{C}.5$[4].

[4]Manuscript *Local Convexity Encoding Fields* under preparation

2

A Scalable Fluctuating Distance Field

2.1 A Part-Based Representation for Tumor Shapes

Tumor growth modeling is extensively studied using theoretical and experimental approaches by a variety of disciplines. While majority of the current studies are focused on modeling microscopic phenomena, mathematical models that operate at a macroscopic level are increasingly investigated through the analysis of clinical medical images [91]. Inhomogeneous and anisotropic tumor growth mechanisms lead to deviations of the tumor's shape characteristics from a compact structure and include protrusions. It is clear that extracting and quantifying the spatial information that irregular tumor shape parts carry would be a helpful macroscopic research tool for a better understanding of the dynamics of tumor growth.

As for clinical usage, the quantification and segmentation of the protruded and peripheral tumor regions could play an important role in radiosurgical applications. The goal of radiosurgery is to deliver a necrotic dose of radiation to the tumor while minimizing the amount of radiation to healthy brain tissues, especially to dose-sensitive tissues [92]. Series of beam configurations are determined as an optimization problem for treatment planning process such that beams will intersect to form a high dose at the tumor ROI. The rapid decrease at the edges of the radiation beam, which corresponds to the between

80% and 20% isodose lines, is called the penumbra region and is generally located on the peripheral regions of the tumor [93]. A model that allows the distinguished analysis of the peripheral regions and segmentation of these parts that receive less radiation dose would not only be useful for isodose planning, but also for evaluating the success of the operation on protrusions and peripheral regions that are in close relation to critical anatomical structures. We propose an interactive method to distinguish protruded-peripheral parts using solely distance relations.

2.1.1 Related Work

Segmentation or partitioning of shapes as boundary meshes is a problem of great interest for geometric modeling and computer graphics fields. The partitioning of the object represented by the mesh into meaningful parts, referred to as *part-type segmentation* by Shamir[15], is highly motivated by the study of human cognition [94, 95]. For an in-detail analysis of existing mesh segmentation methods we refer to [15, 96], along with recent successful approaches [97, 98] and a comparison of part-type segmentation techniques can be found in [99]. Distance functions described on the shape surfaces are commonly utilized for shape decomposition. There is a variety of surface metrics, e.g. geodesic [100], isophotic [101, 102], diffusion [103, 104, 105], volumetric part aware [106]. Though successful with a mesh representation, adaptation of these decomposition methods that use distance metrics to a volumetric representation would not be plausible. Additionally, partitioning the protrusions of tumors would require the abstraction of peripheral regions beforehand, else the association of partitioned boundary segments to the tumor volume would not be possible.

A sound approach for regional shape partitioning is utilizing the medial axis of symmetry, i.e. skeleton representation [64]. Partitioning shapes by associating regions with medial locus branches is very common and also successfully utilized in medical imaging [107, 108, 7, 8]. However, skeletal representations commonly suffer from certain instabilities. One of the instabilities is due to boundary perturbations, which are commonly addressed using smoothing or branch pruning approaches, which involve discarding branches that contribute little to the reconstruction of the shape[109, 110, 57]. For partitioning,

choice of branches to prune would affect the resulting decomposition drastically considering the highly compact shapes of tumors, which also tend to inherit symmetries. Another kind of instability occurs in the regions near the junctions, which is mainly referred to as the ligature problem [111, 112]. A variety of methods have been proposed to cope with the ligature problem, including detecting transitional areas [113], a Bayesian formulation for estimating likely branches that would produce the shape [114] or disconnected skeleton approaches [115, 85, 116]. Additional to these inconsistencies, the association of branches with protrusions is not straightforward and even under slight deformation the abstraction of the centrality of the shape is not possible for fold-symmetry cases, which are highly possible for tumor shapes. Tari's model of Three-Partite-Skeleton, which arises from fluctuating distance fields [117] adresses this problem, which is highly motivating for the purpose of protrusion segmentation.

The fluctuating distance field [1, 117] contains both positive and negative values, and its zero crossing separates central and peripheral volumes. The maximum value of the field can be considered as a rough approximation of the center point for the shape in question, for instance the tumor, whereas the local minima correspond to rough approximations of center points for the protruded parts on the shape. The level curves encode the spatial relationships so explicitly that the separate protruded parts can be segmented even using a watershed segmentation without any additional processing. The extracted central region is compact and the peripheral region is always partitioned, unless it is a perfect annulus. In this model, no control exists over the ratio of region cardinality of positive field values to that of the negative field values. However, such a property can be an advantage in forming a shape field that respects a certain scale of central to peripheral regions of the shape. Particularly for shapes of tumorous structures, where boundaries between peripherality versus centrality is rather vague, variation of such a scale will introduce a flexibility in following shape analysis stages.

2.1.2 Our Contribution

In this paper, we describe a scalable fluctuating distance field as a tumor description model. This model allows the user to interactively adjust the ratio of positive and neg-

ative domain sizes. The corresponding parameter can be set according to nature of the application. Thanks to this addition, a hierarchy of parts is not to be abstracted from the field as in [1]. Instead, fields that represent different hierarchical assumptions are formed, with the trade-off of losing linearity of the formulation. Details about the formulation and implementation of the shape field will be described in Section 2.2, where the fluctuation scale space that arises with the new parameter is introduced and exemplified on 2D shapes and 3D tumor volumes.

The constructed shape fields will be used for an alignment of baseline and follow-up tumor structures. In this registration problem, the distance transform is often used as a shape representation that describes the spatial relationships within the moving and fixed shapes[118]. The adjustment of the location of the zero-level set of the new distance field impairs the effect of scale changes to the resulting field for a fixed fluctuation scale, making the field a robust underlying shape representation for registration purposes. The registration process is described in Section 2.3 and experiments using both synthetic data and patient data are evaluated in Section 2.3.1, where the scalable fluctuating distance representation is compared to the conventional distance transform representation.

2.2 Scalable Fluctuating Distance Field

The concept of fluctuating distance fields, introduced by Tari [1], involves the exploitation of local and global spatial interactions to achieve a field that consists of both negative and positive values. The zero-level set partitions the shape domain into Ω^+ and Ω^-, which corresponds to the central region, a coarse and compact shape, and the peripheral region, which includes all the protrusions of the tumor, respectively. The ridge points on the surface yields the Three-Partite skeletons indicated. Our main motivation in using the fluctuating distance field is the information inherently coded in the resulting level curves at the peripheral regions, which will allow the explicit treatment to peripheral regions for further analysis. In this section we will describe our modification of this method, which will provide the required flexibility and interactivity for our purpose. We will follow by introducing the arising scale-space and illustrating segmented protruded parts using

different fluctuation scales for 2D shapes and 3D tumor volumes.

The fluctuating distance field, $\omega : \Omega \rightarrow \mathbb{R}$ is a real valued function on a discrete lattice, $\Omega \subset \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$, with a neighborhood system, $\mathcal{N}$. ω is generated by the minimization of linear combinations of regional and boundary energies, which are described over the shape domain Ω, as a function of ω.

2.2.1 Energy Terms

The regional energy consists of local and global terms that function as spatial regularizers. Tari [1] proposed a global regional energy, which is the squared average over the domain, connecting all the nodes using a global mean constraint:

$$E_{Global}(\omega_{i,j,k}) = \left(\frac{1}{|\Omega|} \sum_{(l,m,n)\in\Omega} \omega_{l,m,n} \right)^2 \tag{2.1}$$

Differentiating the sum of $E_{Global}(\omega_{i,j,k})$ over Ω leads to the following expression:

$$\frac{\partial E_{Global}(\omega_{i,j,k})}{\partial(\omega_{i,j,k})} = \frac{2}{|\Omega|} \sum_{(l,m,n)\in\Omega} \omega_{l,m,n} \tag{2.2}$$

which would be minimized if ω is composed of all zeros or is a fluctuating function, where positive and negative values cancel each other.

The local regional energy functions as a smoothness term. We use the sum of squared differences between neighboring pixels in a 6 neighborhood system, $\mathcal{N}(i,j,k)$ to obtain the required spatial smoothness for the ω field:

$$E_{Local}(\omega_{i,j,k}) = \sum_{(l,m,n)\in\mathcal{N}(i,j,k)} (\omega_{l,m,n} - \omega_{i,j,k})^2 \tag{2.3}$$

Differentiating this energy w.r.t $\omega_{i,j,k}$ results in the following expression, where Ł corresponds to the seven-point discretization of the laplacian operator:

$$\begin{aligned} \frac{\partial E_{Local}(\omega_{i,j,k})}{\partial(\omega_{i,j,k})} &= -2(\omega_{i+1,j,k}+\omega_{i-1,j,k}+\omega_{i,j+1,k}+\omega_{i,j-1,k}+\omega_{i,j,k+1}+\omega_{i,j,k-1}-6\omega_{i,j,k}) \\ &= -2Ł(\omega_{i,j,k}) \end{aligned} \tag{2.4}$$

The boundary energy is defined for formulating the interactions along the level surfaces. The preservation of interactions between the nodes is imposed on the ω field using the

usual distance transform as a bridge [1]. Thanks to this constraint, central regions of the shape, where the distance transform has larger values have much higher tendency to get positive ω values. The similarity to the distance transform function is formulated as follows:

$$E_{Bdry}(\omega_{i,j,k}) = (\omega_{i,j,k} - \mathcal{D}_{i,j,k})^2 \tag{2.5}$$

where $\mathcal{D}$ denotes the distance transform of the shape. The derivative of E_{Bdry} w.r.t $\omega_{i,j,k}$ is then given as follows:

$$\frac{\partial E_{Bdry}(\omega_{i,j,k})}{\partial(\omega_{i,j,k})} = 2(\omega_{i,j,k} - \mathcal{D}_{i,j,k}) \tag{2.6}$$

Minimization of the combination of these energies results in a ω field that has low expected value, thus fluctuating (2.2), locally smooth (2.4) and resembling the distance transform of the shape (2.6).

2.2.2 A Sign Constraint to Control Fluctuation Scale

The natural location of the zero-level curve under the given constraints often becomes too close to the tumor boundaries, turning out to be a disadvantage while estimating a deformation between two ω fields. In addition, the ability to control the location of the zero crossing turns the ω field to a robust feature for an interactive tool for segmenting the protrusions on the tumor. Therefore we describe an additional global constraint to adjust the position of the zero crossing. The term is constructed as a quadratic expression forcing the sum of the signs of all nodes to be close to a predetermined ratio of the domain size, $|\Omega|$:

$$E_{Sign}(\omega_{i,j,k}) = \left(\left(\sum_{(l,m,n)\in\Omega} sign(\omega_{l,m,n}) \right) - \eta|\Omega| \right)^2 \tag{2.7}$$

where $\eta \in [-1, 1]$ corresponds to the ratio of the intended sum of the signs of all ω points to the number of points in the shape domain $|\Omega|$. While minimizing (2.7), η is chosen as the desired ratio of :

$$\eta = \frac{\sum_{(l,m,n)\in\Omega} sign(\omega_{l,m,n})}{|\Omega|} = \frac{|\Omega^+| - |\Omega^-|}{|\Omega^+| + |\Omega^-|} \tag{2.8}$$

Differentiating the sum of this energy w.r.t. $\omega_{i,j,k}$ would give :

$$\frac{\partial E_{Sign}(\omega_{i,j,k})}{\partial(\omega_{i,j,k})} = 4 \sum_{(i,j,k)\in\Omega} \left(\left(\sum_{(l,m,n)\in\Omega} sign(\omega_{l,m,n}) \right) - \eta|\Omega| \right) \cdot \delta(\omega_{i,j,k}) \tag{2.9}$$

For the approximation of the signum function in a differentiable manner, we used a regularized Heaviside function, then the impulse function $\delta(z)$ was approximated as the derivative of $H(z)$:

$$sign(z) = 2H(z) - 1 \simeq \frac{2}{\pi} \arctan(\frac{z}{\epsilon}), \quad \delta(z) \simeq \frac{1}{\pi}(\frac{1}{1+(\frac{z}{\epsilon})^2})(\frac{1}{\epsilon}) \tag{2.10}$$

where ϵ determines the steepness of the smoothed step and the impulse functions.

A Formulation

The computation of ω is achieved by calculating the steady state solution to the linear combinations of the energy derivatives, which are described above. The combination of the energies is presented in a continuous formulation as follows:

$$\begin{aligned} \iiint_{\Omega} (\omega_{x,y,z} \quad - \quad \mathcal{D}_{x,y,z})^2 + \left(\frac{1}{|\Omega|} \iiint_{\Omega} \omega(\alpha,\beta,\theta)\, d\alpha d\beta d\theta \right) + (\nabla\omega(x,y,z)\,)^2 + ... \\ ... \left(\left(\iiint_{\Omega} sign(\omega_{\alpha,\beta,\theta}) d\alpha d\beta d\theta \right) - \eta|\Omega| \right)^2 dx\, dy\, dz \end{aligned} \tag{2.11}$$

The solution is obtained by applying the method of gradient descent in the following expression:

$$\frac{\partial \omega_{i,j,k}(\tau)}{\partial(\tau)} = -\frac{\partial(\beta_1 E_{Local}(\omega_{i,j,k}) + \beta_2 E_{Global}(\omega_{i,j,k}) + \gamma E_{Sign}(\omega_{i,j,k}) + \beta_3 E_{Bdry}(\omega_{i,j,k}))}{\partial \omega_{i,j,k}}$$

where β and γ values are Lagrange multipliers for the given energies. As natural choices, β_1, β_2, β_3 parameters can be interpreted as 1 [1] . γ is the only Lagrange multiplier that calibrates the relationship between the values of $E_{Bdry}(\omega_{i,j,k})$ and $E_{Sign}(\omega_{i,j,k})$. γ only affects convergence speed when it is within appropriate limits, that is not larger than the maximum value of the $\mathcal{D}$. We choose it as a normalization to the E_{Sign} of the ω field with the desired size of $|\Omega^+|$ using roughly a spherical zero-level set assumption. The iterative

scheme on ω is revealed after an artificial time discretization in τ:

$$\begin{aligned}\frac{\omega_{i,j,k}^{n+1}-\omega_{i,j,k}^{n}}{\Delta\tau} &= Ł(\omega_{i,j,k}) - \frac{1}{|\Omega|}\sum_{(i,j,k)\in\Omega}\omega_{i,j,k}^{n} - (\frac{1}{|\Omega|}\omega_{i,j,k}^{n} - \mathcal{D}_{i,j,k}) \\ &\quad -\gamma\sum_{(i,j,k)\in\Omega}\left(\sum_{(i,j,k)\in\Omega} sign(\omega_{i,j,k}^{n}) - \eta|\Omega|\right)\delta(\omega_{i,j,k}^{n}) \qquad (2.12)\end{aligned}$$

For the third term above, as ω is calculated up to a scale, a weight of $1/|\Omega|$ is used as a weighting between the $\mathcal{D}$ and the ω field.

2.2.3 A Space of Fluctuation Scales

The effect of the parameter η of the E_{Sign} term is not only to change the location of the zero-level set. Its combination with the zero-mean constraint changes the encoding characteristics of the whole domain. For instance, positive η values force the negativity of the nodes that belong to Ω^- much more compared to $\eta = 0$ to satisfy the zero mean constraint. The reason is that there are less number of nodes that are negative, so those have to be more negative to satisfy the zero mean condition. The opposite goes for the negative η values. This causes a diversity in the characteristics of the fields as η changes. A separate normalization can be applied to the positive and negative parts of the fields, which diminishes this effect if not desired.

We depict the resulting fluctiation scale-space for a hand shape in Fig. 2.1(a), where $\omega(\tilde{x}, \eta)$ is presented for $\tilde{x}$ on a vertical line on the hand shape domain and the surface plot for the zero-crossing contour as a function of η is presented in (b) . Notice that the zero-level set sweeps the whole domain smoothly from boundary to central regions, as the information regarding Ω is encoded for different scales of peripherality.

The computed field is shown for three different η values ($> 0, = 0, < 0$) for the symmetric shape silhouette in Fig. 2.2. Note that there are two levels of hierarchy in the peripheral regions of the shape, which can be seen as five different parts at a coarser level, later which are further differentiated into two separate parts. Varying the fluctuation scale parameter, one can capture those two levels of scale (coarser and finer) as can be observed in the resulting field with positive and negative η values, respectively.

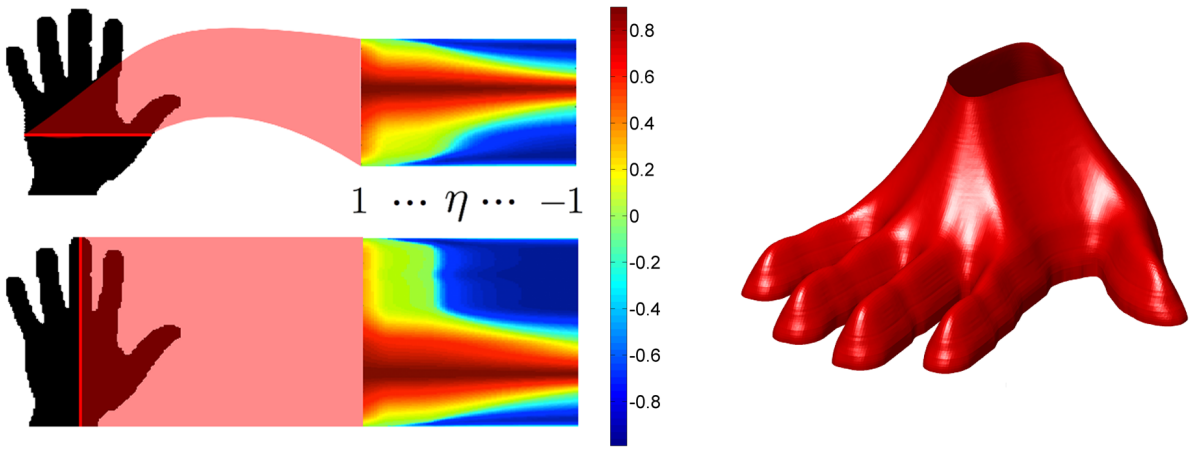


Figure 2.1: *Left* The normalized field $\omega(x = \tilde{x}, \eta)$, where $\tilde{x}$ is shown by horizontal(top) and vertical(bottom) red lines. Image obtained by sweeping η from 1 to -1. *Right* Surface plot for $\omega(x, y, \eta) = 0$

A similar effect is achieved for the leaf silhouette in Fig. 2.3. Using a simple watershed segmentation [119], the resulting partitions reveal the three main leaves with $\eta = 0$, whereas the partitioning with the $\eta > 0$ field reveals the smaller protrusions on those three leaves. Here, the encoding of coarse to fine shape details nicely demonstrates the hierarchical aspect introduced into the fluctuating distance field.

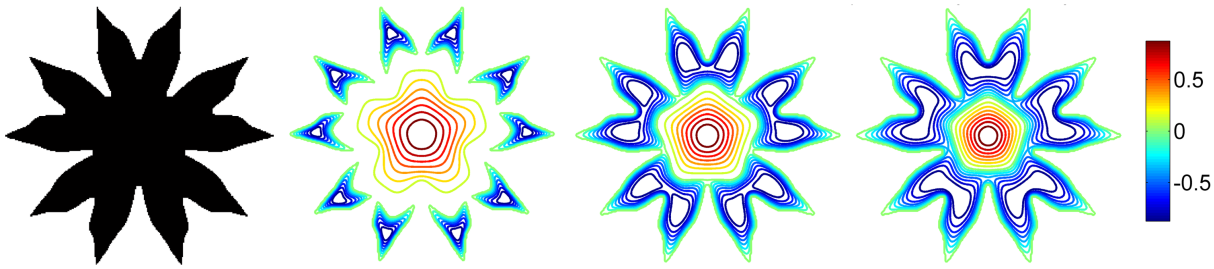


Figure 2.2: *From left to right:* Input shape, ω for $\eta > 0$, ω for $\eta = 0$, ω for $\eta < 0$

We show the original w field and the scalable w field for various η values in Fig. 2.4 for an elephant and a cat silhouette. The first columns next to the silhouettes show the original field followed by the fields with increasing values of the fluctuation scale. The top picture is the whole w field, whereas the lower depicts only its Ω^- partition. Looking at the

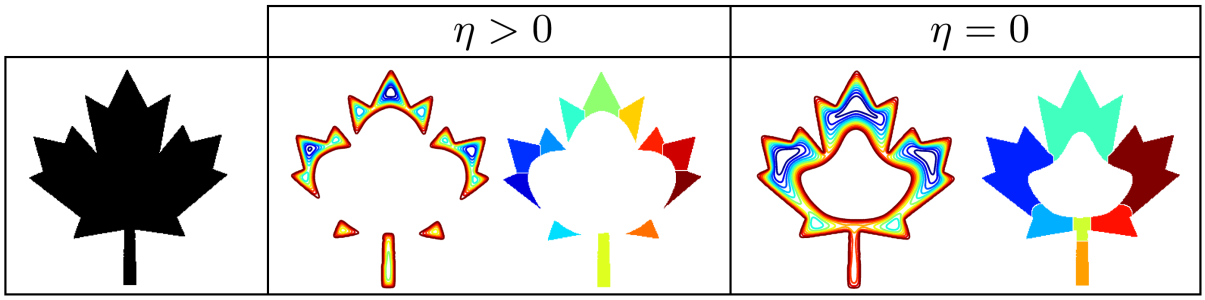


Figure 2.3: Ω^- domain and watershed segmentation results for: *left* $\eta > 0$, *right*: $\eta = 0$

details at the legs of the fields more closely, for instance, the elephant's both front legs are merged in the original w field, as well as for the scalable field for smaller η values. When η is increased (e.g. see the rightmost field), the legs are separated, as can be observed in the Ω^--part of the field. This is because where the two legs are joined, there is a single local maximum with the original and low scale parameter fields, whereas there are two separate local maxima for each leg with the high-scale-parameter field. The same observation holds for the various shape fields over the cat. Note the rear-most leg of the cat and its tail which share a joint single maximum, whereas that extremum separates into two separate maxima for the tail and the rear leg towards the higher η-scale. Another point to remark over these experiments is the interesting feature of the low-η-fields when compared to the original w-field. Note the cat's front legs, and elephant's rear legs, which seem to have a separate maximum for each leg in the original shape field. However, the low η shape fields facilitate to peek at those same features first jointly then separately as the fluctuation scale varies from low to high. As these experiments demonstrate, the hierarchy over the shape is not built from the w-field as in [117], however, we modify the field itself to create the hierarchy that is sought for.

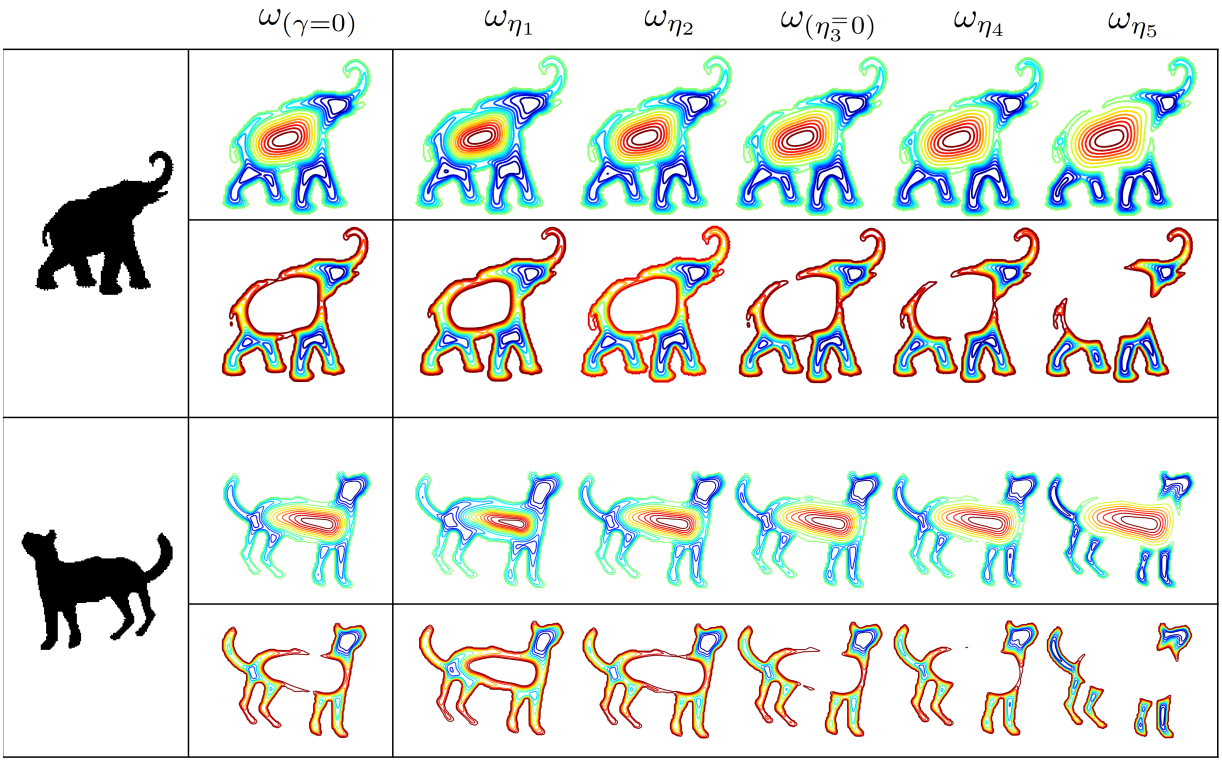


Figure 2.4: The original ω field(left)[1], where the Lagrange multiplier γ is chosen as zero in Eq.2.12 and five ω fields (right) calculated using increasing values for η, where $\eta_1 < \eta_2 < \eta_3 = 0$ and $\eta_5 > \eta_4 > \eta_3 = 0$. Upper row for both shapes is a contour plot of normalized ω and bottom rows depict ω for solely Ω^-.

2.2.4 Interactive Tumor Protrusion Segmentation

The segmentation of the protruded tumor regions is achieved using the information in the negatively-valued regions of the ω field, which encapsulates local minima that depicts separate protrusions. The tumor should be segmented prior to the calculation of ω, for this purpose we use the Tumor-Cut method [120]. A contrast enhanced T1 MRI axial slice is depicted in Fig.2.5, along with the ω field calculated on the tumor shape domain. Partitioning of the negatively-valued domain into protruded parts can be performed using the watershed transform [119] on the Ω^- field. The parts segmented from the resulting ω field can be observed Fig. 2.6 for a sample 3D tumor volume.

With the flexibility that E_{Sign} provides, the size of the positive compact part Ω^+ can be

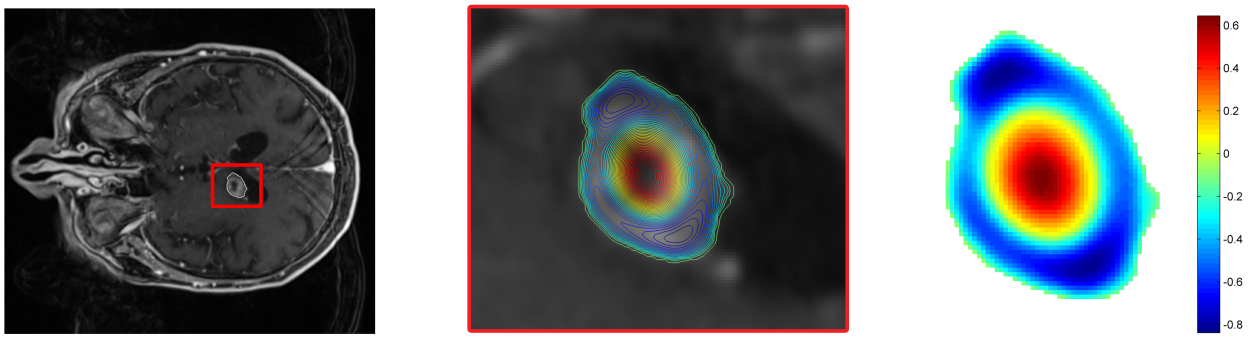


Figure 2.5: *Left:* An axial slice of contrast enhanced T1 MRI of a patient with a tumor. *Middle:* ω field isocontours for the corresponding tumor slice. *Right:* ω field visualized.

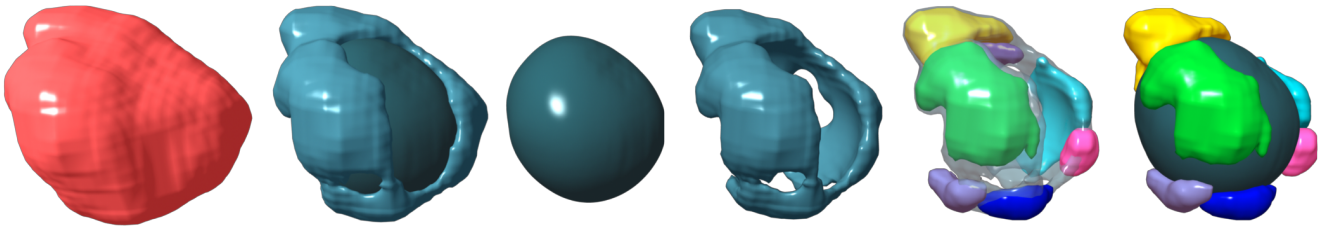


Figure 2.6: *From left to right:* Tumor volume. Positive and negative parts of the proposed field. Positive part of the field. Negative part of the field. Segmented protrusions of the tumor enveloped in the negative part of the field. Segmented protrusions visualized with the positive part of the field.

adjusted with user interference by medical experts or can be calculated automatically by relaxing the η parameter until a predetermined hypothesis regarding the separated volumes are satisfied. The effect of η parameter on the resulting protruded parts is presented in Fig. 2.7.

2.3 Tumor Follow-Up Registration Using ω fields

In order to obtain a valid and unbiased comparison between the performances of ω field and the conventional distance transform $\mathcal{D}$ as underlying shape representations, we chose attributes that are essential in many of the registration algorithms that were proposed

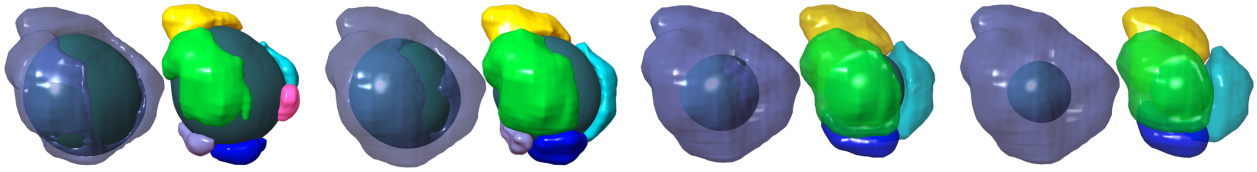


Figure 2.7: Visualizations of positive(opaque) and negative(transparent) parts of the tumor field paired with corresponding segmentation results. The fluctuating distance field for each pair were generated using the corresponding η value.

to calculate such deformations and combine them to end up with a basic yet powerful registration routine.

As linear data terms are not capable of performing well in case of large displacements, we used non-linear data terms and a coarse to fine warping approach which is a well studied combination in the area of optical flow estimation [121]. We follow the traditional model, formulated by means of an energy optimization problem, where deformation is calculated as a mapping between domains of shape fields ω_1 and ω_2. The displacement field $u \in \mathbb{R}^3 = (u_1, u_2, u_3)$ describes the deformation between the tumor and the follow-up shape domains: $u : \Omega_1 \in \mathbb{R}^3 \rightarrow \Omega_2 \in \mathbb{R}^3$. In the following: $x \in \mathbb{R}^3 = (x_1, x_2, x_3)$. The assumption of constancy of the underlying shape representation is formulated as:

$$\omega_1(x) - \omega_2(x + u) = 0$$

In addition to this data term, a regularization term based on the gradient of the deformation field is utilized. Following the original Horn and Schunck optical flow model [122], the combined functional F, where α is a parameter that controls the smoothness term :

$$F(u) = \int_{\Omega_1} (\omega_1(x) - \omega_2(x + u))^2 + \alpha^2(|\nabla u_1|^2 + |\nabla u_2|^2 + |\nabla u_3|^2)\mathrm{d}x \tag{2.13}$$

is minimized to yield the Euler-Lagrange equations, which are non-linear due to the $\omega_2(x+u)$ terms they contain. The first order Taylor expansions are used for those terms to obtain the linear system of three equations. First one of those three equations (for each coordinate) is written as:

$$(\omega_1(x) - \omega_2(x + u) - \nabla\omega_2(x + u)\, du)\, \omega_{2_{x_1}} + \alpha^2 div(\nabla u_1) = 0 \tag{2.14}$$

where $\omega_{2_{x_i}}$ is the spatial derivative of ω_2 w.r.t x_i and $du \in \mathbb{R}^3$ describes an unknown update to the known variable u. In its solution, we adopted the warping scheme, introduced in [121], where the deformation field u is set to zero at the coarsest level and updated by $u^{n+1} = u^n + du$, as soon as du is computed at each finer scale using an inner loop of SOR iterations. $\omega_2(x + u)$ is computed at the beginning of each outer iteration by applying a warping process to $\omega_2(x)$ using the deformation field, u^n. The number of outer iterations depends on the downsampling factor. In order to achieve the full potential of the model, instead of the conventional 0.5 downsampling factor, we used a fixed value of 0.95 and a large number of outer iterations.

2.3.1 Registration Results and Discussion

Using both 2D synthetic shapes and real patient 3D tumor volumes, the performance of the ω field as an underlying shape representation for tumor follow-up registration is demonstrated against the distance transform, which is the conventional method to impose spatial shape relationships to the registration procedure.

Synthetic data results for pre-smoothed distance transform, pre-smoothed normalized distance trasform and fluctuating distance fields are respectively demonstrated in Fig. 2.8. On the top row for each of the experiments, the white and gray shapes denote the fixed and moving objects respectively, where the displacement vector field is demonstrated using arrows. Local volume change for each of the corresponding displacement field is generated using the determinant of the deformation gradient ($det(I + \nabla_x u)$) and presented below. The values of the determinant that are greater than 1 indicate a local expansion, whereas values less than 1 indicate a local contraction.

The distance transform, $\mathcal{D}$ is invariant to rotation and translation, but it is quite sensitive to scale changes[118]. Without a normalization, $\mathcal{D}$ representation can perform well for deformations without scale changes only, which certainly is not the case for tumor followup analysis. On the other hand, normalization causes an ambiguity in the information preserved in $\mathcal{D}$, leading to an estimation of the deformation field that does not fully describe the change between the shapes. However the ω field adopts less ambiguity, since the information is partitioned to separate parts, which leads to a robust estimation

of the deformation. E_{Sign} constraint contributes highly to this robustness to scale change, for the ratio η will be the same in Ω_1 and Ω_2. Our experiments are highly coherent with this description. In Fig.2.8, it is clear that $\mathcal{D}$ without a normalization fails to produce a smooth vector field. In addition while the local volume change in the deformation fields estimated using ω is in accordance with the change in the shapes, the normalized $\mathcal{D}$ representation approaches fail to generate intuitive results. The expansions and contractions at the peripheral regions in Fig. 2.8 reveal the counter-intuitive nature of the displacement vectors generated using normalized $\mathcal{D}$.

Our experiments with patient data are demonstrated in Fig. 2.9, where the estimated 3D vector fields are visualized (on the left) for two pairs of tumor volumes on each row. Those pairs of tumor volumes are obtained after a segmentation on a pre-therapy and follow-up MRI scan and undergo a large change in terms of global scale. In addition, we present the displacement fields to a specific protrusion (Fig. 2.9 in the middle), which was segmented as described in Section 2.2.4. Various 2D cross sections are also depicted on the right along with the local volume change maps using deformation gradient determinants as explained above. Considering the large motion, necessity of regularization is quite larger in 3D tumor volumes compared to the phantom data in Fig. 2.8. For that reason the differences in the volume change maps are not as distinctive for the 3D volumes. But when these subtle changes are analyzed, they reveal the strength of the ω field in contrast to $\mathcal{D}$. The volume change maps of the second tumor shape in Fig. 2.9 is a convincing example: When the upper slice is analyzed it is clear that the deformation calculated using ω field (on the right) describes the compression smoother, yet on the bottom slice (right), it successfully represents the expansion while the distance transform approach is too smooth to describe an expansion. A similar robust behavior can be observed on the given local volume changes of the first tumor pair in Fig. 2.9.

2.4 Conclusions

Motivated from physical significance and clinical relevance in follow-up problems, we proposed a method to analyze the protruded and peripheral regions of tumor shapes. In

order to introduce a parameter to control the fluctuation scales, we modified the fluctuating distance field [1] with an additional constraint on the ratio of sizes of the positive and negative domains which indicate central and peripheral shape regions respectively. This modification led to an interactive framework for segmenting the protrusions and partitioning tumorous structures, albeit the loss of the linearity of the original shape field model. The introduced nonlinear term due to its variable scale parameter, i.e. the "fluctuation scale", facilitates a hierarchical encoding of parts of the shape silhouette. By varying the fluctuation scale from low to high values, it is possible to observe the coarse to fine levels of hierarchy both in the field and its segmentations even by utilizing a very simple segmentation method.

The scalable shape field becomes a potentially powerful underlying shape representation for shape registration procedures, due to an increased robustness to scale changes without losing the information it inherits particularly in terms of the parts of a shape. For the registration application, the representation performance of the field was demonstrated in comparison to the conventional distance transform by observation of local volume changes in a tumor follow-up problem. Some counter-intuitive local changes were obtained by the latter, while the expected expansion and compression properties between pre-therapy and follow-up tumor volumes were provided by the deformation field estimated between the part-based shape fields.

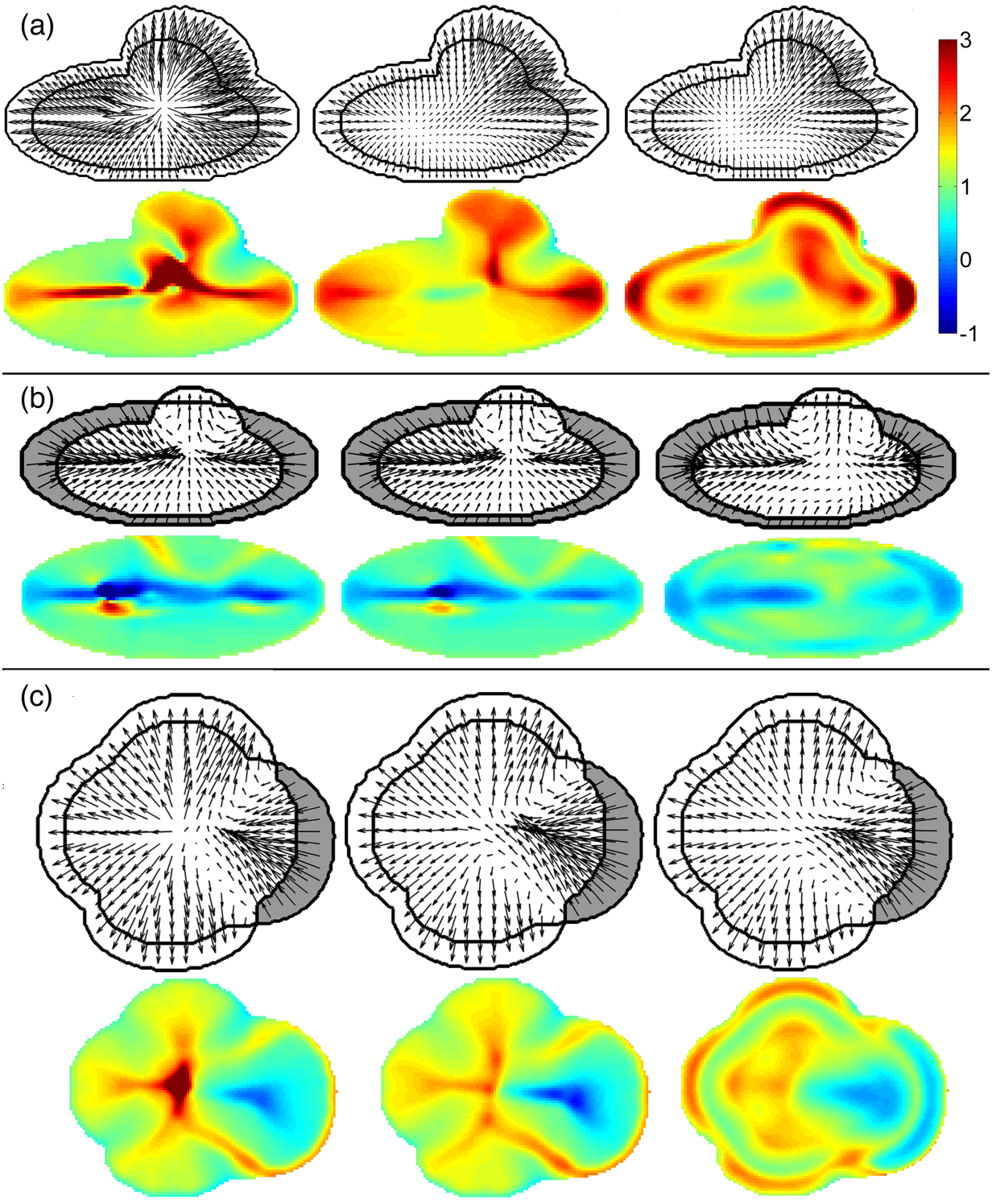


Figure 2.8: (a,b,c): Visualizations of deformation field vectors and volume change pairs for registration of each synthetic shape couples, generated using Left:Distance Transforms Middle: Normalized Distance Transforms. Right: Scalable Fluctuating Distance Fields.

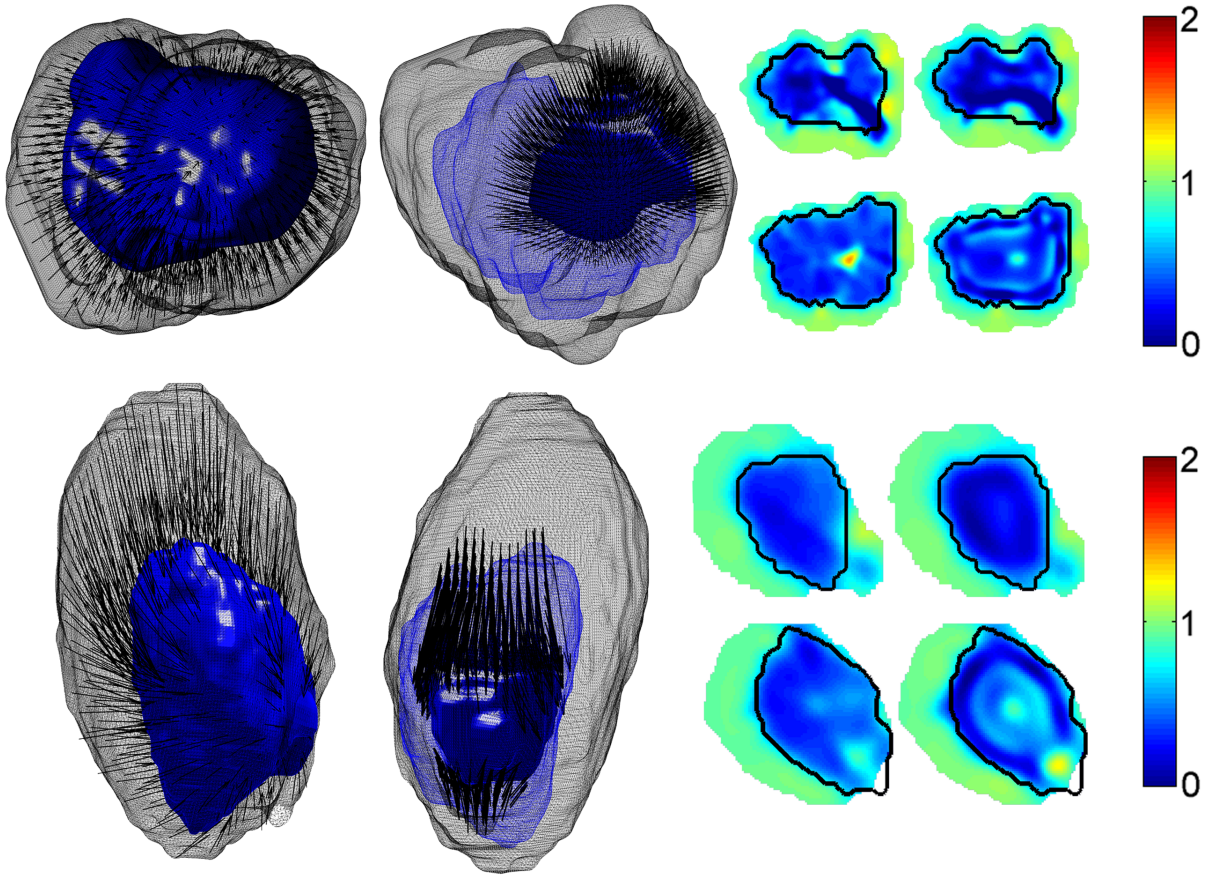


Figure 2.9: For both parts of the figure: *Left:* Displacement field vectors from gray initial tumor to blue followup tumor. *Middle:* The displacement vectors to a specific segmented protrusion. *Right:* Local volume change maps in initial tumor domain for selected axial slices of the tumor shapes, the black contours denote the followup tumor. The maps on the left and right are generated from the deformation fields calculated using normalized $\mathcal{D}$ and ω fields respectively.

3

Screened Poisson Hyper-Fields

3.1 Introduction

A novel perspective on shape characterization is presented using the screened Poisson equation. We discuss that the effect of the screening parameter is a change of measure of the underlying metric space; also indicating a conditioned random walker biased by the choice of measure. A continuum of shape fields is created, by varying the screening parameter or equivalently the bias of the random walker. In addition to creating a regional encoding of the diffusion with a different bias, we further break down the influence of boundary interactions by considering a number of independent random walks, each emanating from a certain boundary point, and the superposition of which yields the screened Poisson field. Probing the screened Poisson equation from these two complementary perspectives leads to a high-dimensional hyper-field: a rich characterization of the shape that encodes global, local, interior and boundary interactions.

3.1.1 Related Works

In the late 90's and the following decade, elliptic PDEs started to appear as alternative models for computing smooth distance fields. In [85], screened Poisson PDE is employed:

$$\begin{aligned} \Delta v - \frac{v}{\rho^2} &= 0 \\ v|_{\partial\Omega} &= 1, \end{aligned} \tag{3.1}$$

where $\frac{1}{\rho^2}$ is the screening parameter that controls the level of smoothing. The approximate distance field created by this PDE is smooth and differentiable, and has smooth level sets, in contrast to the level sets of the distance transform obtained from the Eikonal equation. With any given ρ value, the field's value of 1 at the shape boundary drop towards the interior of the shape. While a motivation in [85] was to create a shape scale space, demonstrated particularly for shape skeletons via the controlled smoothing parameter, in [52], the intuition of a random walker starting at an interior point and its mean hitting time required to reach the shape boundary led from its discrete interpretation to the continuous Poisson equation with zero Dirichlet boundary conditions on the shape boundary. Various measures based on the solution field were extracted and shape properties were used for classification of shapes as well as actions [123]. In [124], the authors also utilized the Poisson equation to derive a shape characteristic measure based on the variation over the streamlines of the solution field, and used it to differentiate between the shapes of anatomical structures for healthy and diseased populations. Recently, Poisson equation is revisited as a tool for robust skeletonization [125, 126]. In [127, 128], a connection between nonlinear Hamilton-Jacobi equations, for which the Eikonal equation is a special case, and the screened Poisson equation by taking $\rho \longrightarrow 0$ is presented, along with an efficient approximate distance transform computation using FFT. The importance of the linearity of shape embedding space was brought into attention by the work of [129], [130] that represented contours as zero level set of a harmonic function in the solution of the Laplace PDE. The linearity property, which was also emphasized in [127], enables proper addition of shape fields, facilitating creation of shape template or atlas representations that stay within the original spaces of shapes. [103] solved heat flow with a fixed time parameter and used its normalized gradient field to obtain the closest scalar potential field with the same gradient. In [131], smooth distance fields are considered as L_p distance fields, where p is the control variable. A recent shape field related to the screened Poisson [1] is a fluctuating one consisting of both negative and positive values inside the shape by addition of a zero-mean constraint to the shape field. The zero-level set then partitioned the shape domain into two: one that corresponds to the central region, a coarse and compact shape, and one to the peripheral region, which included protrusions from a shape.

The screened Poisson PDE was employed for several other applications with a typically fixed screening parameter: for image processing applications as in image filtering and sharpening of [132]; for mesh filtering applications as in anisotropic and interactive geometric filtering over meshes of [133]; and for surface reconstruction in [134]. [132] started from a variational perspective by writing out the gradient of an unknown function to be close to a given vector field as well as a term of data fidelity to a given function which "screens" the 2D Poisson equation. This was then Fourier transformed to show that the screened Poisson can be interpreted in frequency domain as a filtering operation for images, while it can be solved using an FFT or DCT. [133] extended [132] to meshes for localised editing by changing the Riemannian metric of the underlying space, proportional to surface curvature, as well as a multi-grid implementation of the equation. The effect of the fidelity value, i.e. the screening parameter, was also discussed to result in more dampening and amplification at low frequencies with smaller parameter values. [134] modified this method by putting positional constraints, i.e. the data fidelity, only over a set of input points rather than over the full domain. Adding a screening term to the Poisson surface reconstruction framework, the screening parameter was also adjusted to the resolution in a multi-grid implementation.

In a parallel line of research, from the heat equation perspective, the multi-scale property of the heat kernel led to development of shape signatures that take advantage of heat diffusion process on surfaces [135]. This line of work makes use of the spectral properties of the Laplace-Beltrami operator, which is the generalization of the Laplace operator from the Euclidean space to a Riemannian manifold. In [135], the heat kernel signature (HKS) at a point on the shape manifold is defined in terms of the weighted sum of the squares of the eigenfunctions at the point. The weights are given by the exponentials of the negated eigenvalues multiplied by the temporal variable t in heat flow. It was shown that under certain conditions (*i.e.*, if the eigenvalues of the operator are not repeated) the heat kernel signature is as informative as the family of heat kernel functions parameterized both in space and time. The HKS also relates to global point signatures [136], which are based on eigenfunctions normalized by square root of the corresponding eigenvalues, and to diffusion distance [137, 105, 138] between two points over the shape manifold, which is

defined by the distances between the eigenfunctions at those two points. [139] constructs a scale-invariant HKS (SI-HKS) by logarithmically sampling the time-scale that translates into a time shift, which is then removed through taking a Fourier transform modulus to overcome the scale sensitivity of HKS. A volumetric extension of HKS was shown in [140].

Recently, Wave Kernel Signature [141] based on complex Schrodinger equation is presented as an alternative to HKS. The authors make the point that HKS employs a collection of low-pass filters parameterized by time variable, causing the suppression of high frequency shape information whereas the WKS captures both the high and the low frequency shape information.

Meanwhile, works such as Shape DNA [59] showed the utility of the eigenvalues of the Laplace operator, where the distances between shapes were expressed as the p-norm of the difference between the truncated eigenvalue sequences for the two shapes. In [58], a normalized shape DNA distance, called the weighted spectral distance is proposed.

Laplace-Beltrami eigenfunctions of surfaces proved to be extremely useful in applications of 3D shape matching and retrieval. In [142], it was shown that a bijective mapping between a given pair of shapes induces a transformation of a function of derived quantities between them. Furthermore, this transformation can be written as a linear map between selected basis functions over both surfaces, exemplified by the Laplace-Beltrami eigenfunctions. [143] presented a method to perform shape matching in a reduced space in which the symmetries of shapes were identified and factored out. This was achieved within the functional maps framework of [142] where the functional linear map was decomposed into its symmetric subspace and its orthogonal subspace, and the former was utilised to carry out the shape matching between symmetric shapes. For joint analysis of multiple shapes, [144] presented a coupled construction of common Laplacian eigenfunctions using approximate joint diagonalizations.

In [145], a shape-aware interior-mesh distance was defined by propagating a distance measure defined on the mesh to the surface interior, while preserving distance properties. This was exemplified by the diffusion distance and mean-value coordinates selected as the barycentric coordinates. [146] later applied this idea to interpolating the Laplace-Beltrami eigenfunctions of the boundary into the interior volume by using barycentric coordinates.

This way, a volumetric measure was constructed from the HKS, i.e. the interior HKS, and adopted to finding correspondences between volumes and shape retrieval.

3.1.2 Our Contribution

Both the Poisson and the screened Poisson equations found increased utility in various shape descriptors. As the screening parameter in (3.1) tends to ∞, the screened Poisson equation approaches to the Poisson equation. The controlled smoothing provided by the screening parameter is advocated by some researchers and recent works [132, 126, 128, 1, 103, 134] rejuvenated the model.

Our work differs in several aspects. We consider multiple instances of the screened Poisson equation to decompose the sources of variability due to several factors including the boundary sources and the screening parameter, both of which are novel. We discuss that the effect of screening parameter is a change of measure of the underlying metric space, hence, fixing ρ^2 fixes the measure. Suitably sampling N values for the screening parameter and m points for the shape boundary $\partial\Omega$, we form a stack of $N \times m$ screened Poisson fields. We call this collection as a screened Poisson hyper field. This is not a scale-space in the usual sense but hides in it a two-dimensional scale space of shapes, coarsening in the direction of increasing ρ^2 and decreasing field values. We argue that the hyper field is a full characterization of all sorts of interactions between shape elements: local-global and boundary-interior. Then we discuss two low-dimensional embedding schemes, one to unveil parts and the other to produce consistent mappings, which we call Screened Poisson Encoding Maps (SPEM), for the purpose of shape matching and shape retrieval.

Encoding a change in the diffusion using the varying screening rates in the screened Poisson equation forms a remarkable parallelism with the class of methods in spectral shape analysis. We argue that a coverage of ρ^2 parameter space for (3.1) over the shape domain brings advantages over the coverage of the temporal parameter space for the heat kernel over the shape in terms of producing a direct volumetric shape representation. [140], extending the heat kernel signatures to volumes, noted that the boundary isometries of the HKS do not carry over to volume isometries, however, volumetric HKS can still faithfully model nearly isometric deformations, which are argued to be more suitable in

modelling of natural articulations and deformations of solid objects. On the other hand, [146] propagated HKS on the surface towards the interior of the shape to be able to construct volumetric measures to benefit from nice properties of the HKS, including the surface isometry, the multi-scaleness and insensitivity to topological noise at small scales however at the expense of its sensitivity to scale of the shape [140, 147]. Different from these earlier heat-kernel based approaches, here, we directly compute volumetric distances from the solution to the volumetric screened Poisson PDE, which enjoys properties such as multi-scaleness based on a varying screening-parameter that tunes smoothness of the level curves of the field, an adaptation to scale by an appropriate mapping, and a near isometry-invariance as demonstrated experimentally by the robustness of the proposed method under a 3D nonrigid shape retrieval application (§ 3.5.5).

As an alternative to the heat equation and its kernel, our work presents a different differential operator, a different kernel, and demonstrates the high-ranking performance of the SPEM to articulated pose and deformation in a publicly available large-scale benchmark data set: *SHREC'11 Shape Retrieval on Non-Rigid 3D Watertight Meshes* [87]. Our method, as shown in the presented 3D shape retrieval application, provides a robust and high-performance alternative to those methods based on shape's intrinsic surface properties. Furthermore, existence of fast solvers for the screened Poisson PDE, as realised by [132, 133, 128, 134] in other applications of image filtering and mesh processing, is another factor that makes it attractive to be adopted in a new shape representation idea as in this paper.

The organization of the paper is as follows. In § 3.2, we show separation of the sources of variability in the v-field, and present the construction of the new shape hyper-field. We expound properties of the new hyper-field and the SPEM in § 3.3 through a random walk interpretation, relation to geodesic distances, and a connection to spectral methods. In § 3.4 we present how decompositions on shape hyper fields via two alternative techniques produce consistent mappings and part decompositions. Finally in § 3.5, we present our experimental results followed by conclusions.

3.2 A new hyper-field

In this section, first the existing two-dimensional scale space parameterized by ρ and the values of v (§ 3.2.1) is explained. Then we describe the two dimensions of the new shape representation: the varying of the ρ (§ 3.2.2), and decomposition of the boundary sources (§ 3.2.3). The new hyper-field thus includes two dimensions of variability: (i) by variation of ρ, it covers the internal smoothing characteristics of v; (ii) by variation of boundary sources, it covers interactions between individual boundary nodes versus all internal nodes. We note that the decomposition into those two dimensions do not create a true scale-space per se, however, creates a rich shape hyper-field representation from which descriptive volumetric shape encoding maps (SPEM) can be extracted.

3.2.1 A two-dimensional Scale Space

The information encoded in the resulting field v, as a shape representation, is highly dependent on the value of ρ^2. The influence of the parameter ρ^2 can be observed in Fig. 3.1, where different fields that arise using different ρ^2 values are presented for a cat shape. Smaller ρ^2 values lead to fields where distinct relations in the regions that are close to shape boundary (protrusions, indentations) are extracted, but are clueless about the central part of the shape and global interactions. In contrast, larger values of ρ^2 generate fields that are coarse in the regions close to the boundary, but able to capture global interactions within the shape. Unlike the level curves of the solution of the Eikonal Equation, the level



Figure 3.1: v fields for different values of ρ^2.

curves of v (the solution of screened Poisson Equation) has smooth level sets, and as one moves along the gradient lines, the level curves gets smoother. As discussed in [148, 85]

$$v(\boldsymbol{x}) \approx \rho\left(1 + \frac{\rho}{2}\, curv(\boldsymbol{x})\right)\frac{\partial v}{\partial n} + O\left(\rho^3\right) \tag{3.2}$$

where $curv(\boldsymbol{x})$ is the curvature of the level curve of v passing through the point $\boldsymbol{x}$ at $\boldsymbol{x}$, and n is the direction of the normal. Thus, one can imagine a two-parameter family of level curves parameterized by v and ρ. Smoothing increases with a decrease in v and an increase in ρ^2. This is a very interesting property. This explains how the linear screened Poisson mimics a non-linear reaction-diffusion. Though this observation was made in the early work of [85], the follow up work on screened Poisson typically focused on isolated treatment of the ρ^2. Rangarajan [128, 127] took a very small value to approximate the Eikonal Equation, while Tari [126, 1] and Shah [149] used very large values.

We believe that isolated treatment is hindering full utilization of the controlled smoothing offered by the model. As we show in §3.4.2, once the entire scale space is utilized, both local and global interactions can be realized and a natural hierarchical central to peripheral decomposition of the shape domain is achieved without requiring the recent non-local term in [1].

3.2.2 Varying ρ^2: Sweeping Internal Smoothing Characteristics

In a setting where the screening parameter is considered as an additional dimension to the spatial ones, it is clear that the $n+1$-Dimensional field calculated for a shape embedded in $\mathbb{R}^n$, where the parameter ρ^2 is swept from 0 to ∞ inherits all the information that is possible to be extracted using such a method about the shape. The collection of fields $\{v^\rho\}_{\rho^2}$ consists of a 1D family of functions that sweeps the ρ^2 dimension for each node on the lattice that the shape is described on. A field created using only one of these values would explain only limited portion of the variance. In order to capture this high dimensional information, we linearly sample ρ^2 values to N bins, and calculate v^ρ for each ρ_j^2 value for $j = 1, ..., N$. Each v^ρ field as a single instance explain relatively little variation of the shape in comparison to the whole family.

We depict via an example that the v function is coding characteristics that extend beyond the distance to the nearest boundary point as well as curvature (Fig 3.2). We consider several nodes in a shape domain. They are marked with colored crosses. Each node has a different character: The blue one is central; the other four are closer to shape boundary, pink being the closest; and red is on a peripheral part (finger). The v versus ρ

plot on the right depicts striking differences among $v(\cdot, \rho)$ profiles for these different shape nodes. For example, the two points colored red and pink respectively have closer profiles as they have comparable proximity to the shape boundary. However, the profiles are not nearly identical because the red node is residing in a thin part of the shape while on the contrary the pink one is not.

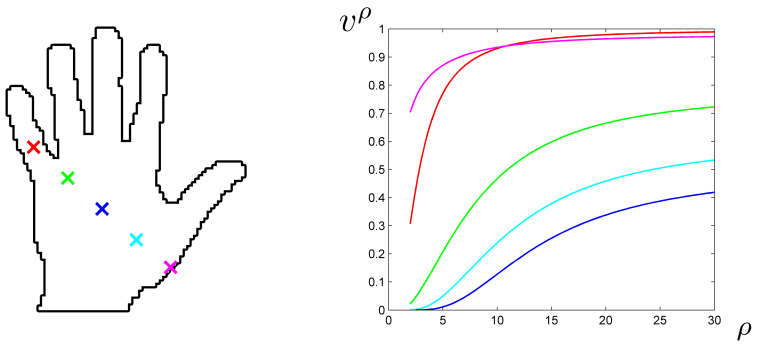


Figure 3.2: Field value versus ρ at five selected nodes of distinct characters. v function is coding characteristics that extend beyond usual distances. A dense linear sampling is used between $\rho = 2$ and $\rho = 30$.

Fig. 3.3 demonstrates further coding characteristics of the v^ρ field. A set of 1D profiles $(v^\rho(\tilde{\boldsymbol{x}}))$ for a set of locations $\tilde{\boldsymbol{x}}$ on a hand silhouette are depicted. Here, the point we emphasise is that the selected locations, $\tilde{\boldsymbol{x}} \in \Omega$ are equidistant to the shape boundary. Observe that the 1D curve describing the relation between v and ρ shows a quite different character for each point, which has the same Euclidean distance to the boundary, while the v-field is able to encode the diversity of the geometric shape information among those points.

3.2.3 Fixing ρ^2: Decomposition of Boundary Sources

The two-dimensional scale space is a continuous collection of simple closed curves parameterized by $[1, 0) \times \{1, 2, \cdots, N\}$. For a fixed screening parameter, a one-dimensional scale space is formed by the collection of the level curves of the field v^ρ, which is a union of these level curves. This is not the only way to envision v^ρ. Thanks to the linearity

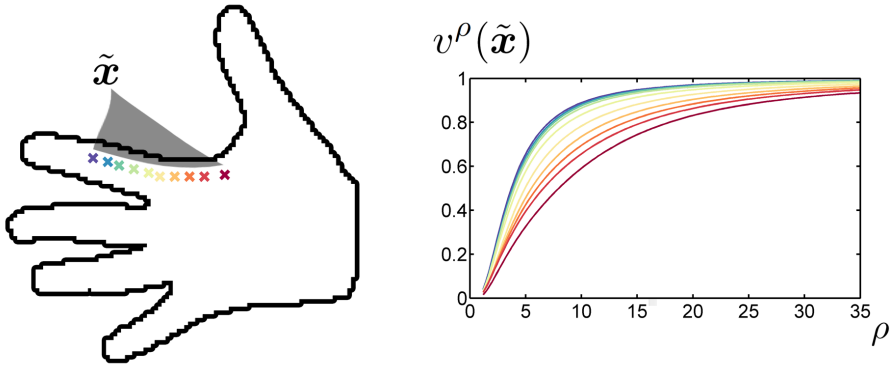


Figure 3.3: Behavior of v^{ρ} in ρ dimension for sampled points on the domain that are equidistant to the boundary.

of the equation, it is also possible to express v^{ρ} as a superposition of basis fields each of which is expressing the contribution due to a single "unit" of inhomogeneity.

In order to elaborate on the super-positioning aspect of the screened Poisson PDE for a fixed ρ^2 and better understand geometric properties induced by boundary interactions, we consider decomposing the sources of inhomogeneity in the boundary condition. Assuming that the shape boundary is given as a set of points $\partial\Omega = \{\boldsymbol{p_1}, \boldsymbol{p_2}, ..\boldsymbol{p_m}\}$, we consider m independent PDEs:

$$\begin{aligned} \Delta v^{\boldsymbol{p_i}}(\boldsymbol{x}) - \frac{v^{\boldsymbol{p_i}}(\boldsymbol{x})}{\rho^2} &= 0 \\ v^{\boldsymbol{p_i}}(\boldsymbol{p})|_{\boldsymbol{p}\in\partial\Omega} &= \delta(\boldsymbol{p} - \boldsymbol{p_i}) \end{aligned} \tag{3.3}$$

where $v^{\boldsymbol{p_i}}$ denotes the solution when the only inhomogeneity is due to the point $\boldsymbol{p_i} \in \partial\Omega$. Thanks to the linearity of the equation, these "sub-fields" are the building blocks that make up the field v described in (3.1):

$$v = \sum_{i=1}^{m} v^{\boldsymbol{p_i}} \tag{3.4}$$

The super-positioning of the sources is demonstrated on a 1D example in Fig.3.4. The boundary condition on the third column is an addition of the two boundary conditions used in the first two columns. Hence the solutions in the third column are superpositions of the pair of solutions given on the respective row of the first two columns.

In Fig. 3.5, the logarithm of the field $v^{\boldsymbol{p_i}}$ obtained from a boundary point $\boldsymbol{p_i}$ on the hand shape is visualized on the left. It can be observed that the v field shows a sharp fall

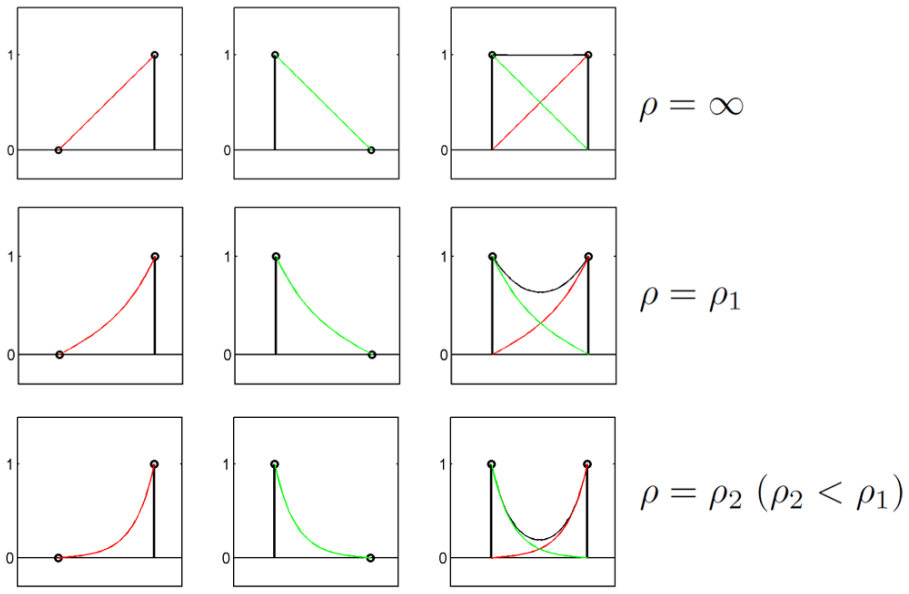


Figure 3.4: Solutions of the screened Poisson equation for a 1D experiment using three different boundary conditions (columns) and three different ρ values (rows).

of its values over the fingers whereas a much less steep slope of fall is observed from the boundary points of the hand's side palm regions (e.g. close to the wrist). This different behaviour is expected. To analyse it on a simpler case, assume a spherical geometry with a source term $s(\boldsymbol{r})$ at the origin, and consider the Poisson equation: $\Delta v = s(\boldsymbol{r})$, the fundamental solution is: $g(\boldsymbol{r}) \propto \frac{1}{\boldsymbol{r}}$, whereas for the screened Poisson equation: $\Delta v - \frac{v}{\rho^2} = s(\boldsymbol{r})$, the fundamental solution reads: $g(\boldsymbol{r}) \propto \frac{e^{-\frac{\boldsymbol{r}}{\rho^2}}}{\boldsymbol{r}}$ [150]. Hence with a nonzero source term the solution is given by:

$$v(\boldsymbol{r}) = \int_{\Omega} d\boldsymbol{r}' s(\boldsymbol{r}) \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} e^{-\frac{|\boldsymbol{r}-\boldsymbol{r}'|}{\rho^2}}. \tag{3.5}$$

For a spherical symmetric case, the source is diffused to its surrounding points by a convolution with a kernel inversely proportional to the distance between the source and the given point for the standard Poisson case, whereas for the screened Poisson the convolving kernel is in addition weighted by a decaying exponential. Although for the arbitrary geometric configuration of our boundary conditions we cannot write an integral equation to solve for the result, we can observe the exponential decay effect in our v-field from a single source point to other points. With a union of all boundary sources, the effect

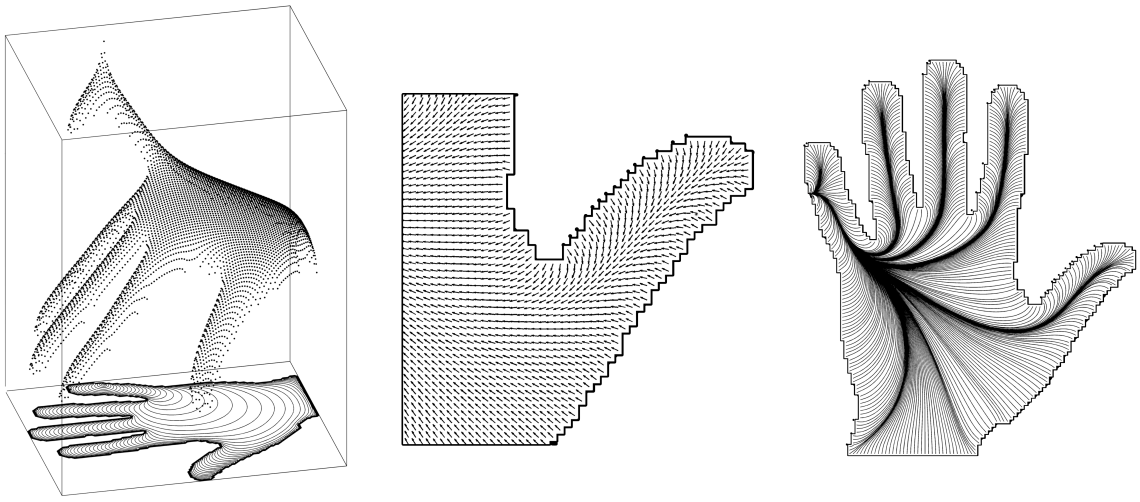


Figure 3.5: Restricting the boundary inhomogeneity to a single point $\boldsymbol{p_i}$ on the little finger. a) Iso-contours (bottom) and values of v-field using $log(v^{\boldsymbol{p_i}})$ visualised as a point cloud; b) Normalized gradient, $\frac{\nabla v^{\boldsymbol{p_i}}}{|\nabla v^{\boldsymbol{p_i}}|}$ for the 'thumb'; c) Streamlines obtained by tracking along the normalized gradient directions.

is even more pronounced. Similarly in § 3.2.2, we changed the rate of decay by varying ρ to probe this property. We will further discuss the relation of the v-field to geodesic distances in § 3.3.2.

3.2.4 Putting it altogether: The New Hyper-field

By considering a total of $N \times m$ screened Poisson equations, we form a stack of fields. This stack of fields hides separation of several sources of variability due to all kind of interactions: local, global, region and boundary. The schematic depiction is given in Fig. 3.6. Intuitively, this can be best explained as simultaneous decomposition layers.

In the first decomposition layer, ρ^2 is varied to obtain a stack of fields $\{v^{\rho_i}\}_{i=1\cdots N}$. Each slice in the stack is an interpretation of the shape with a certain bias – choice of measure, and is a collection of shape boundaries embedded as level curves hence parameterizable by a continuous parameter $s \in (0, 1]$. This is the second layer of decomposition. The stack of fields $\{v^{\rho_i}\}_{i=1\cdots N}$ as parameterized by $(0, 1] \times \{1, \cdots, N\}$ defines a 2D scale space of shapes, coarsening in the direction of increasing ρ^2 and decreasing s. At the final decomposition

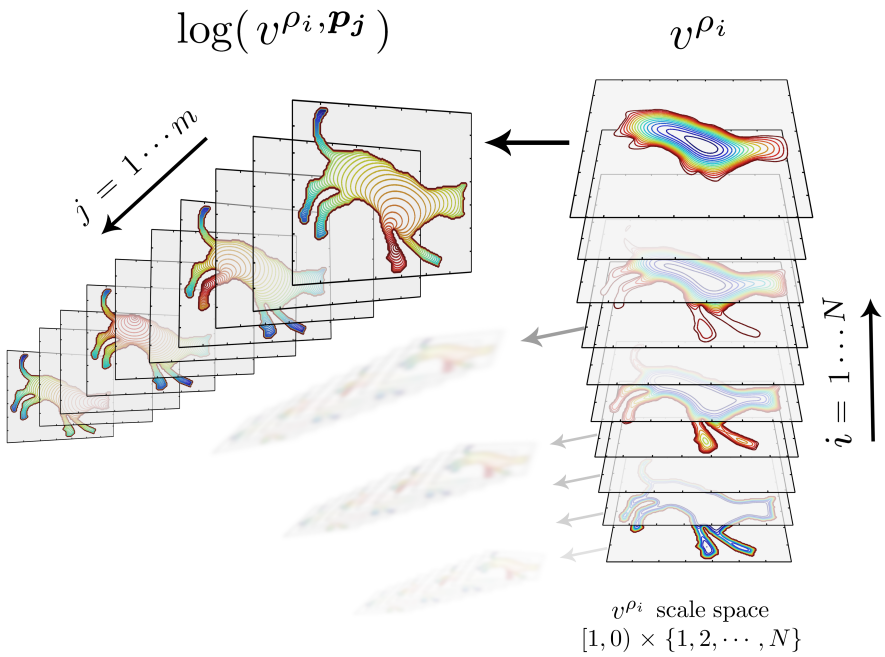


Figure 3.6: Separating sources of variability in the shape hyper-field.

layer, the effect of inhomogeneity (note that the solution to the PDE in (3.1) is the trivial solution in the absence of inhomogeneities) is individuated by considering m-fields $v^{\rho_i, \boldsymbol{p_j}}$ for $j = 1 \cdots m$, which add up to $v^{\rho_i} = \sum_{j=1}^{m} v^{\rho_i, \boldsymbol{p_j}}$. This last layer added by the boundary source sweep is built on top of the nonlinear scale-space of ρ and level curves of v, hence maintains a more complex structure.

The hyper field provides a rich characterization of the shape. We will present how to extract this information in a robust way in §3.4.

3.3 Screened Poisson: Properties

3.3.1 Screened Poisson as a conditioned random walk

In this section, we expound the underlying stochastic interpretation of the v^ρ field in order to gain more intuition into its coding properties. Specifically, we are interested in understanding better the effects of 1) the change of ρ, and 2) the boundary interactions.

First, let us shift the inhomogeneity in the boundary condition in (3.1) to the right hand side as a source term, and then consider an inhomogeneous heat equation: $(\Delta +$

$\frac{\partial}{\partial t})u(\boldsymbol{x}\,,t) = f(\boldsymbol{x}\,)$. On one hand, the steady state solution as $t \to \infty$ is $1 - v^{\rho}$ for $\rho \to \infty$ (*i.e.*, the solution of the Poisson Equation). On the other hand, the transient solution is

$$u(\boldsymbol{x}\,,t) = \int p_t(\boldsymbol{x}\,,\boldsymbol{y}\,)f(\boldsymbol{y}\,)d\mu(\boldsymbol{y}\,) \tag{3.6}$$

where μ is the Lebesgue measure and $p_t(\boldsymbol{x}\,,\boldsymbol{y}\,)$ is the transition probability from point $\boldsymbol{x}$ to $\boldsymbol{y}$ in time t. The transition probability (also called heat kernel) is given by the Gaussian function:

$$p_t(\boldsymbol{x}\,,\boldsymbol{y}\,) = \frac{1}{(4\pi t)/2}\exp\left(-\frac{|x-y|^2}{4t}\right) \tag{3.7}$$

Now let μ be a measure on a Riemanian manifold $\mathcal{M}$. The inhomogeneous heat equation with the corresponding Laplace (-Beltrami) operator on the manifold is

$$(\Delta_{\mu} + \frac{\partial}{\partial t})u(\boldsymbol{x}\,,t) = f(\boldsymbol{x}\,)$$

The transient solution is given by (3.6). Let us examine the effect of screening following Grigoryan [151]. We note that introducing screening to the Poisson equation corresponds to a change of measure. Let $\widetilde{\mu}$ be the new measure, then $\Delta_{\widetilde{\mu}}$ is related to $(\Delta_{\mu} - \frac{1}{\rho^2})$ by the Doob h-transform.

$$\Delta_{\widetilde{\mu}} = \frac{1}{h}\circ(\Delta_{\mu} - \frac{1}{\rho^2})\circ h \longrightarrow \Delta_{\widetilde{\mu}}v = \frac{1}{h}(\Delta_{\mu} - \frac{1}{\rho^2})(vh) \tag{3.8}$$

To summarize, diffusion is a stochastic Markov process, indeed a Brownian motion with heat kernel as its transition probability. In the case of the diffusion governed by screened Poisson, the new transition kernel $\widetilde{p}_t$ that relates to the original transition kernel is the heat kernel on the Riemannian manifold with measure $d\widetilde{\mu} = h^2 d\mu$ [151]. For a random walk on a network, when $p_t(\boldsymbol{x}\,,\boldsymbol{y}\,)$ is induced by conductances c_{xy}, then $\widetilde{p}_t$ is induced by conductances $\widetilde{c}_{xy} = h(x)h(y)c_{xy}$ [152] [153]. This means that the conditioned random walk behaves like the unconditioned walk but is biased by an isotropic drift h.

The conditioned random walk with a certain ρ^2-value affects a point in the shape domain with a certain bias, making it possible to probe multiple random walkers going through different conductances over the shape. We believe that this is how the continuum of fields encodes the shape characteristics both locally and globally with its varying screening rates or biases. This can also be interpreted as Brownian motions with different

drift amounts, the zero drift corresponding to the unconditioned random walk, hence pure diffusion without any screening term of the standard Poisson equation.

For a fixed ρ^2, the field $1-v^\rho$ is a superposition of multiple random walks on a manifold with a measure $\widetilde{\mu}$. Note that the transient solution (after the change of measure) for the time-dependent equation then would be given by

$$1 - v^\rho(\boldsymbol{x}\,, t) = h_\rho(\boldsymbol{x}\,) \int \widetilde{p}_t(\boldsymbol{x}\,, \boldsymbol{y}\,) f(\boldsymbol{y}\,) h_\rho(\boldsymbol{y}\,) d\mu(\boldsymbol{y}\,) \tag{3.9}$$

At the steady-state, the transition kernel becomes only a function of distance independent of t. Thus, separating the boundary condition to a set of points, and solving the screened Poisson PDE for each single point as in Eq. (3.3), each field value $v^{\boldsymbol{p_i}}(\boldsymbol{x}\,)$ (after a normalization) is interpreted as the probability that the biased random walker emanating from $\boldsymbol{p_i}$ to arrive at the locations $\boldsymbol{x}$. We note that the intuition of the boundary condition on a random walker was mentioned by [103] for the heat flow, with the zero Dirichlet boundary condition implying absorption of heat that leads to random walker "falling off" the grid. With this interpretation, the way we set the point source on the boundary to unity while setting all other boundary points to zero implies that the probability of the walker falling off the grid differs substantially for different local geometric regions of the shape (see Fig. 3.3 for this effect).

3.3.2 Relation to Geodesic Distances

There is a strong link between the values of the $v^{\boldsymbol{p_i}}$ field and the geodesic distance from $\boldsymbol{p_i}$ to another shape node, with the underlying Riemannian metric. A prominent aspect that forms this link is the gradient directions of $v^{\boldsymbol{p_i}}$, which are parallel to geodesics. The choice of boundary conditions configures the resulting gradient field. For instance, Dirichlet boundary conditions attract the flux to the medial locus. In Fig. 3.5 (middle and right figures), we show normalized gradient directions along with streamlines obtained by tracking points in the gradient direction. The link between the heat flow kernel (i.e. the $\rho \to \infty$ case) and geodesic distances was established by Varadhan: $-\sqrt{-4t\ log(p_t)}$, where t corresponds to the amount of time that passes after heat diffusion starts [154]. Simply taking the logarithm of the v field leads to an encoding of the local relationships

in a rather useful manner and preserves the gradient directions. The choice of logarithm stems from the exponential decay of the field (Eq. 3.5) also noted by [85], [127], and the logarithm of the field values become strictly negative, decreasing as the probabilities for the random walkers get less likely. We note that this is not an attempt to make the v-field values similar to Euclidean distances. Taking the square-root as in Varadhan's formula [154] also preserves the gradient directions but suppresses high rates of decay. This sort of treatment would compromise a very desirable property for part based analysis of shapes: at nodes that belong to articulated regions on the shape domain, as the probabilities for random walkers to go off the grid increases, the rate of decay increases drastically. This property was observed in Fig. 3.5 (on the left). Therefore, with the v-field, we are exploiting an exponential decay effect with a complementary contribution from the shape boundary conditions, to construct a beneficial "geodesic distance" from the given shape geometry. Observe the effect of this complementary contribution in Fig. 3.3, where the points that have equal Euclidean distances to the boundary have v^ρ field values which encode a geodesic distance that both shows an exponential character and is more global in the sense that it is affected by the full shape boundary conditions.

3.3.3 Relation to Spectral Methods

The popularity of the heat-kernel-based methods in non-rigid shape matching is due to the usefulness of the heat kernel function in finding near-isometric correspondences between shapes. This is appealing because many expected deformations between shape surfaces, particularly the articulated motion, can be approximated by an isometric mapping. Because the isometry of a manifold preserves the heat kernel [151], heat kernel signature was shown to be isometrically-invariant in [135]. However, a volumetric isometric invariance was not sought for in the volumetric HKS of [140], and it was argued that the articulations and non-rigid deformations of solid objects do not follow a boundary isometry. Similarly, although we do not show an isometry property for our volumetric Screened Poisson Encoding Maps (SPEM), we discuss our approach against the heat-kernel-based approaches next. With μ as the Lebesgue measure, the heat kernel in (3.7)

can be expressed as [151]:

$$p_t(\boldsymbol{x}\,,\boldsymbol{y}\,) = \sum_{k=1}^{\infty} e^{-\lambda_k t}\varphi_k(\boldsymbol{x}\,)\varphi_k(\boldsymbol{y}\,) \tag{3.10}$$

where φ_k are the eigenvectors and λ are the eigenvalues of the Laplace operator: $\Delta_\mu\varphi + \lambda\varphi = 0$. Based on the heat kernel, Ovsjanikov [155] defined the heat kernel map $\Theta_q(\boldsymbol{x}\,) = p_t(\boldsymbol{q}\,,\boldsymbol{x}\,)$, which measures the amount of heat transferred from a source point $\boldsymbol{q}$ to other points $\boldsymbol{x}$ over a given shape surface. The idea is to match the point from the target surface whose heat kernel map is closest to that of the given point in the reference surface. Hence, a correspondence between the two shapes can be established. On the other hand, by varying the t parameter, the heat kernel signature (HKS) [135] creates a 1-parameter family of functions from the diagonal of the heat kernel, also called the auto-diffusivity function: $p_{t_1}(\boldsymbol{x}\,,\boldsymbol{x}\,), ..., p_{t_n}(\boldsymbol{x}\,,\boldsymbol{x}\,))$.

The constructed 1-parameter family of functions based on time t in the heat kernel approaches is similar in spirit to our method. However, rather than the time variable t, we vary ρ variable in the screened Poisson operator. In the former, the temporal evolution of the heat operator is considered, hence the multi-scale heat diffusion characteristic in time is taken into account, whereas in our approach, the 1-parameter family of solution fields to the screened Poisson PDE with varying screening parameters provides the biased diffusion of the boundary sources, from the boundary towards the shape interior. Similar to the Heat Kernel Map [155], it would be possible to match shapes by sampling a set of source points $\boldsymbol{q_j}$ inside the shape and using directly the 1-parameter family of the screened Poisson hyper-fields $\{v_{i=1\cdots N}^{\rho_i,\boldsymbol{q_j}}\}$ at points $\boldsymbol{x}$ on the shape surface. Our work differs by the following: (i) in contrary to the heat kernel map approach, we put the sources on the boundary and diffuse those towards the inside of the shape with a different differential operator, i.e. the screened Poisson; (ii) instead of directly using the 1-parameter screened Poisson fields, we create a low-dimensional embedding of these functions over the ρ-dimension (§3.4.2). The embedding unveils the diffusion bias in projection maps which provide beneficial properties like scale adaptation, compactness and representation power, which are experimentally verified (§3.5.4,3.5.5).

3.4 Extracting information from Hyper-fields

3.4.1 Unveiling parts from the hyper field via sparse coding

We first focus on a single slice of the hyper-field (a fixed measure). This is a collection of m fields that forms a vector field and contains individuated boundary-internal node interactions of the shape. One may construct different useful measures from these interaction vectors. For instance, analysis of correlation between two internal nodes either inside the shape domain or between two boundary nodes, or between a boundary and an internal node are all possible using this collection. Even basic clustering methods such as k-means or Gaussian mixture models will lead to intuitive clusters of internal nodes. However, we chose to employ a specific matrix factorization technique (non negative sparse coding) to portray the decomposability of the global-local information to unveil the parts of shape.

In order to decompose the collection onto a set of components, we start with a normalized *log* field which has zero mean at each point:

$$\begin{aligned} \boldsymbol{V}^{\boldsymbol{p_i}}(\boldsymbol{x}) &= \log(v^{\boldsymbol{p_i}}(\boldsymbol{x})) - \frac{1}{m}\sum_{j=1}^{m} log(v^{\boldsymbol{p_j}}(x)), \\ &= \log(v^{\boldsymbol{p_i}}(\boldsymbol{x})) - \log \sqrt[m]{v^{\boldsymbol{p_1}} v^{\boldsymbol{p_2}} ... v^{\boldsymbol{p_m}}}(x) \end{aligned} \tag{3.11}$$

Note that centering the *log*-field by its mean is equivalent to centering the field by its geometric mean:

$$\boldsymbol{V}(\boldsymbol{x}) = \log\{\frac{\boldsymbol{v}}{\sqrt[m]{v^{\boldsymbol{p_1}} v^{\boldsymbol{p_2}} ... v^{\boldsymbol{p_m}}}}\}(\boldsymbol{x}) \tag{3.12}$$

In order to apply the non-negative matrix factorization, the vector elements that are lower than the mean are replaced by zeros. Though this normalization procedure ignores a region within a certain proximity to the boundary node of interest, thanks to the centering of the data, remaining regions are encoded in a manner that allows distinction of prominent parts. The resulting non-negative vector field can now be analyzed as an additive combination of some bases, leading to a part-based representation. The non-negative measurements obtained by a normalization with the mean and median are depicted in Fig.3.7.

Arranging the measurements $\boldsymbol{V}(\boldsymbol{x})$ into columns of a matrix $Y^{m\times|\Omega|}$ for each shape node $\boldsymbol{x}$, would allow the linear decomposition of the data as $Y \approx AS$, where the matrix A

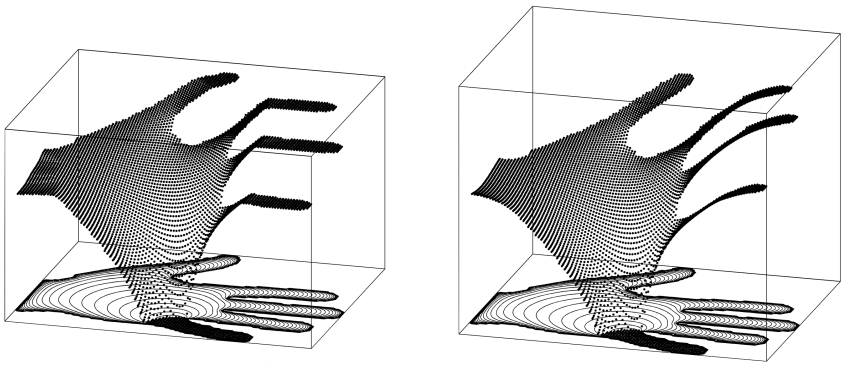


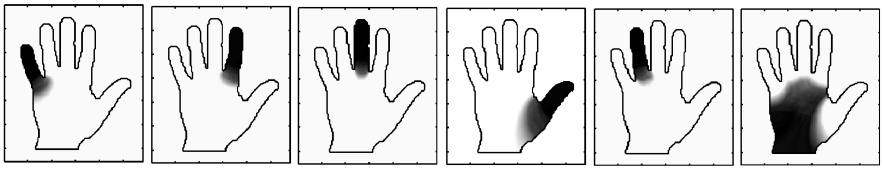
Figure 3.7: Non-negative measurements: $y_j(i)$, where the same $\boldsymbol{p_i}$ in Fig. 3.5 is used. *Left:* Normalization by *median*. *Right:* Normalization by *mean*.

is the mixing matrix with *basis vectors* as its columns. The rows of S contains the hidden components that encodes the contribution of each mixing vector while reconstructing the input vectors. When both factors A and S are forced to be non-negative, the decomposition corresponds to the method of non-negative matrix factorization (NMF) [156][157]. The non-negativity of the factors makes the representation additive as desired.

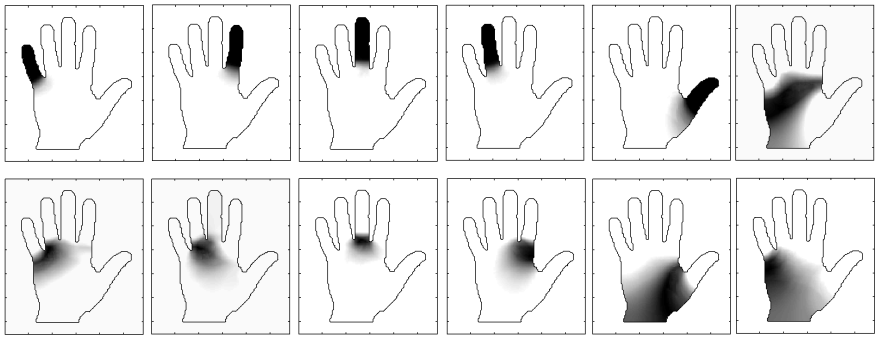
Many variants of NMF have been proposed since the pioneering work of Lee and Seung [157]. Due to its additive nature, NMF produces a sparse representation of the data, where the data is represented using inherent *active components*. Non-negative sparse coding (NNSC), introduced by Hoyer[158], forms an analogy between NMF and sparse coding [159]. NNSC provides control over the sparseness of the hidden components by adding a sparsity-inducing penalty to the objective function, which is very desirable feature for obtaining shape parts as active components that are described on the shape domain. By selecting k mixture elements, the objective function of the NNSC is formulated as:

$$\min_{A\in\mathbb{R}^{m\times k},S\in\mathbb{R}^{k\times|\Omega|}} \sum_{i=1}^{|\Omega|} \left(\frac{1}{2}\|y_i - AS_i\|_2^2 + \lambda\|S_i\|_1\right) \quad s.t.\ A \geq 0,\ \forall i,\ S_i \geq 0, \tag{3.13}$$

where the first term forces minimization of reconstruction error and the second term forces the sparseness. λ controls the tradeoff between sparseness and accurate reconstruction of Y. Sparsity is enforced by using the L_1 norm, this formulation can also be considered as the constraint in the *Lasso* problem [160]. $\lambda = 0$ case is equivalent to NMF formulation (i.e. no additional sparsity). The problem is solved using the method of Mairal et al.



(a)



(b)

Figure 3.8: a) NNSC components obtained using a large λ and $k = 5$; b) NNSC components obtained using a low λ and $k = 12$.

[161][162], which outperforms method of Hoyer [158] in minimizing the objective function in batch mode.

The resulting shape decomposition as NNSC components for a hand shape is presented in Fig.3.8. The results are produced using the slice corresponding to $\rho \to \infty$. In the first experiment (Fig.3.8.a), the shape is decomposed into six components, with a larger λ value. The fingers and the central part of the hand are separated as expected. In the second experiment (Fig.3.8.b), twelve components are obtained with a relaxed sparseness constraint. Notice that the components that represent fingers, which are the most prominent parts, are preserved. Additional components represent the connection points of articulated parts to the central part of hand. Also, the central part of the hand is partitioned into three different parts. The fact that important parts are preserved even when the separation settings are relaxed illustrates the nature of the information preserved in the measurements and robustness of the representation.

3.4.2 Producing consistent mappings for shape correspondence: SPEM

In the previous section, we have concentrated on a single slice in the hyper-field and demonstrated that sparse coding unveils parts by integrating local-global interactions. In this section, we focus on a complementary problem in shape analysis: defining real-valued functions on a shape domain that can be used for the purpose of matching or registration. In the other dimension of the hyper-field, the 1-parameter space that is spanned by varying ρ-values encodes the boundary-interior diffusion characteristics. Although it is possible to utilize this 1-parameter field directly for shape matching, we take one step further and we compactly code the variation in the ρ dimension to produce consistent mappings through a low-dimensional embedding. There is a vast number of dimensionality reduction approaches. We advocate use of principal component analysis (PCA) which produces consistent maps that exhibit adaptation to scale (see Fig. 3.9 and Fig. 3.10).

We find the linear PCA very intuitive as compared to some other popular decomposition methods. Orthogonality of the bases provides quite a consistent mapping across shapes. Independent component analysis based methods form inconsistent mappings. We have observed that non-linear methods such as locally linear embedding [163] or Diffusion Maps [105] over-learn the ρ-space, leading to less number of features and less consistency. Linear PCA also outperforms latent variable methods such as Probabilistic PCA solved by maximum likelihood estimation [164]. The data is created by a *linear* operator, and it is extremely *smooth*. We have observed consistency (among different poses of similar shapes) using a direct singular value decomposition even for the projections that explain variance as small as (10^{-14}).

Consider the $|\Omega| \times N$ matrix Y. Each column of Y is a v field ($v : \Omega \longrightarrow \mathbb{R}$) for a certain choice of ρ^2 and is normalized by the mean of the field in ρ dimension. The covariance matrix of Y is computed, and then decomposed to yield an orthogonal set of bases: the eigen maps $\Phi_k, k = 1, ...N$ of the field. After the new bases are calculated, the field can be projected to form N mappings, where each mapping $\mathcal{P}_k$ is related to a

measure of the variance explained by the k^{th} basis:

$$Y_{|\Omega|\times N} = \begin{bmatrix} v^{\rho_1}(x_1) & v^{\rho_2}(x_1) & \dots & v^{\rho_N}(x_1) \\ v^{\rho_1}(x_2) & v^{\rho_2}(x_2) & \dots & v^{\rho_N}(x_2) \\ \vdots & \vdots & & \vdots \\ v^{\rho_1}(x_{|\Omega|}) & v^{\rho_2}(x_{|\Omega|}) & \dots & v^{\rho_N}(x_{|\Omega|}) \end{bmatrix} - \begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N} v^{\rho_i}(x_1) & \dots \\ \frac{1}{N}\sum_{i=1}^{N} v^{\rho_i}(x_2) & \dots \\ \vdots & \\ \frac{1}{N}\sum_{i=1}^{N} v^{\rho_i}(x_{|\Omega|}) & \dots \end{bmatrix}$$

$$Y^T Y = \Phi\Lambda\Phi, \qquad \mathcal{P}_k = Y\Phi_k \tag{3.14}$$

The low-dimensional embedding facilitates a principled selection of a handful of projections maps, as we call them the SPEM (Screened Poisson Encoding Maps). We note that we observe some interesting properties such as almost perfect representation of the variability in just several bases (or projections). We relate this to the linearity and smoothness of the screened Poisson operator. Using these bases, we observe visual correspondence even in 2D shapes under a perspective transform.

3.4.2.1 Adaptation to Scale

The resulting eigenvectors Φ_n for a hand shape can be observed in Fig. 3.9. The eigenvectors adapt to global changes of the shape, leading to a robust representation. This is exemplified by scaling the hand shape. Notice that the eigenvectors change because a specific characteristic that is detected for a larger distance corresponds to a larger ρ value. This adaptation does not mean that the field is scale invariant, because discretization in spatial and ρ domains would not allow direct invariance. However, the representation does not change abruptly as scale increases. In order to show this, we computed peak-SNR (PSNR) values (in dB) between the original hand shape (maximum value of distance transform is 20 pixels) and its scaled versions up to scale 4 (Fig. 3.9 scale changes are coded by colour on the bottom right). Note the slow monotonical change across scale for the projections, which provides coherence against shape scale changes.

Adaptation of the principal directions in ρ-space to *scale* is also presented in Fig. 3.10. Class of $\{n/4\}$ regular star polygons for $n = 9, .., 20$ are depicted, where all the vertices are lying on circles of a constant radius. As n increases, the shapes become more circular. This change of internal distance relationships affects the characteristics of the

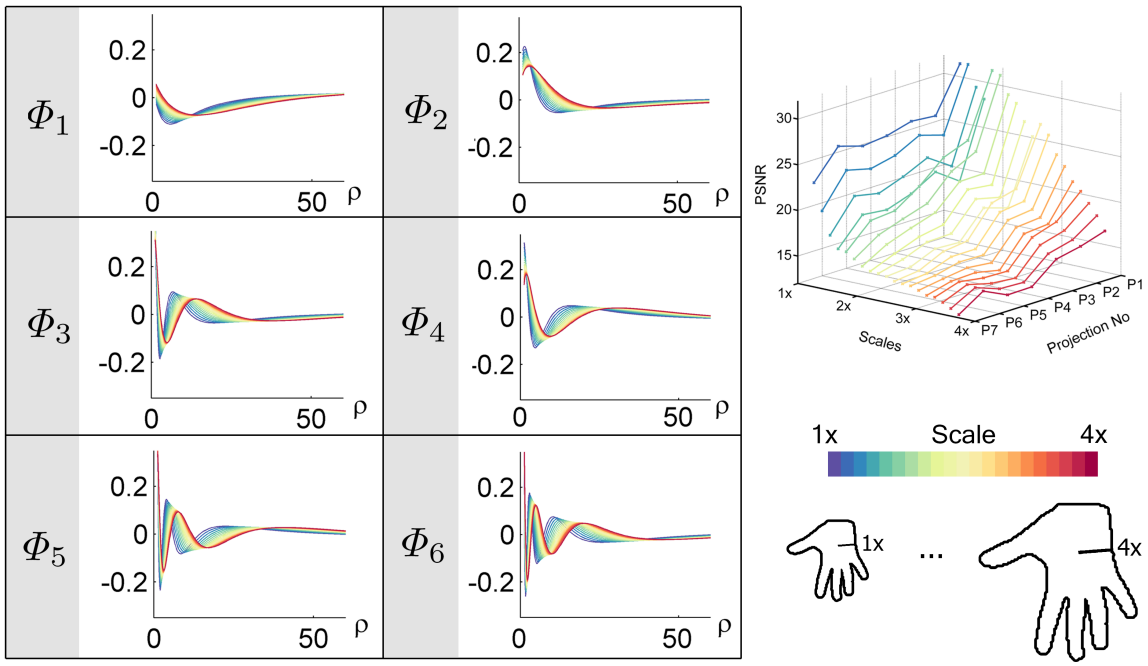


Figure 3.9: For the hand shape: Left: calculated Φ_n colored according to corresponding scale ratios; Right: PSNR values for projections obtained using Φ_n across different shape scales show a slow monotonically changing behaviour, which provides a desired robustness to scale changes, color coded as shown on the bottom right.

hyper-field in its ρ-dimension. The eigenvectors of the covariance matrices are altered in accordance, leading to robustness to scale changes. The six eigenvectors on the right are almost identical and they are calculated for the shapes that are scaled to have the same maximum distance to boundary. This property of the projections imply that the discriminability of the projections originate from local and global spatial relationships. The model offers a framework where globally similar shapes should have similar projections in locally similar regions, which makes it a promising tool for shape analysis along with robustness to global scale effects.

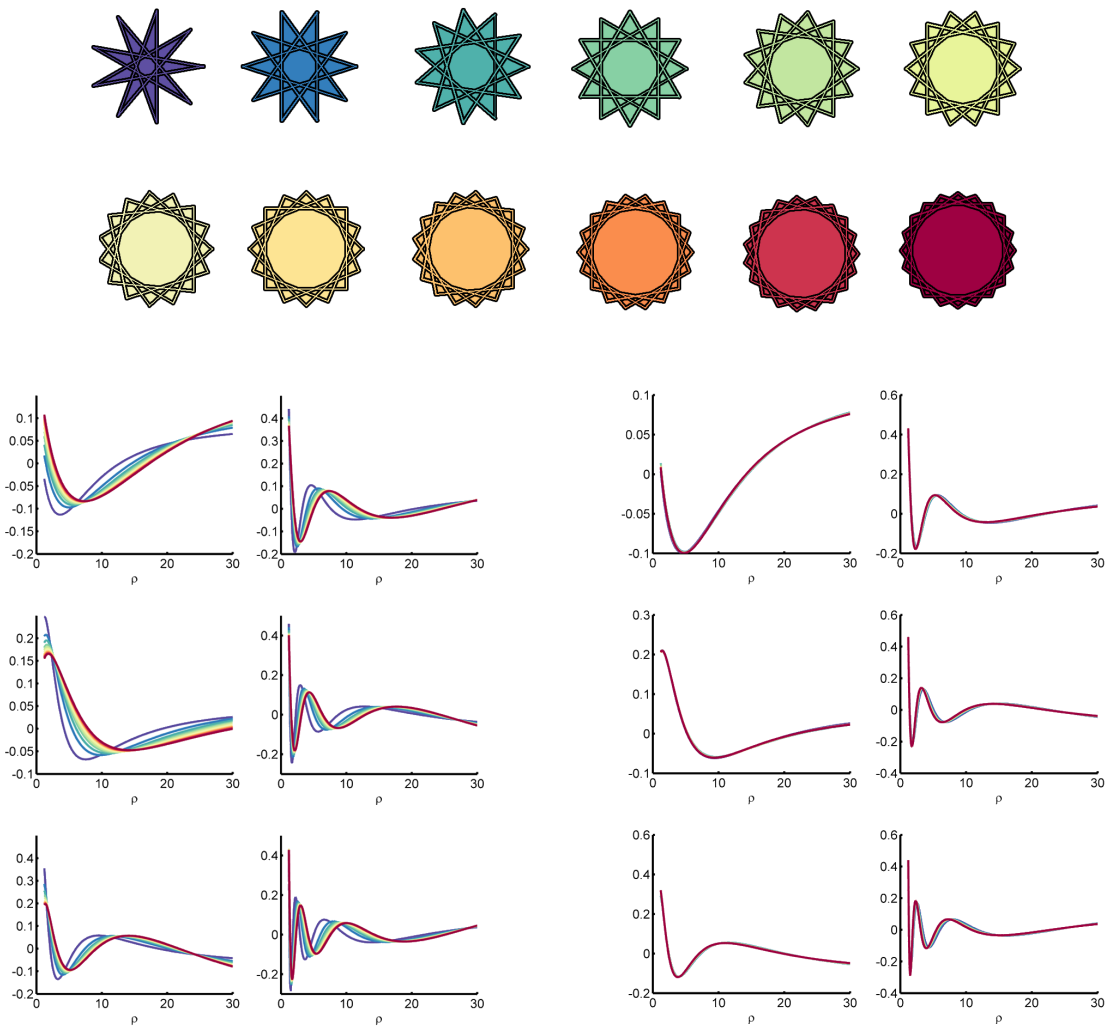


Figure 3.10: *Top:*$\{n/4\}$regular star polygons, for $n = 9, .., 20$ *Left:* First six eigenvectors $\Phi_1, ..\Phi_6$ for the shapes colored accordingly. *Right:* The first six eigenvectors $\Phi_1, ..\Phi_6$ for the shapes after re-scaling with respect to the maximum value of the shortest distance to the boundary.

3.5 Results and Discussions

In this section, we demonstrate the expressive power and robustness of projections of the new hyper-field. After discussing computational issues, we first present qualitative results with sparse coding over boundary decomposition of the shape hyper field. Next, we show first a study on 2D shape classification that demonstrates the usefulness of the SPEM with a moment-feature based evaluation. Finally, we validate the new SPEM descriptor over the 3D *SHREC* benchmark data set [87].

3.5.1 Computational Aspects

Computation of each field v_ρ or $v_\rho^{\boldsymbol{p}}$ can be done in parallel in both approaches that we presented. Notice that we calculate projections $\mathcal{P}_i$ using all the boundary nodes used as sources. Also, we fix ρ for the calculation of the fields for the sparse coding (NNSC) application. Calculating a field for each boundary node taken as a source for a 3D shape is not feasible, yet it is possible to apply a similar approach to calculation of fields over segmented regions on the boundary. This requires a fast initial partitioning of the boundary nodes with large granularity.

Each field is a solution to an elliptic linear PDE, which is a problem that occurs in various fields, and many fast alternative solvers exists. Adaptation of GPU's is an ongoing study for more than a decade [165]. While sparse Cholesky decomposition and FFT based approaches work in subquadratic time [132, 103], multifrontal methods [166, 167] and multigrid methods [52, 168] can reach O(N), which is the lower bound for the problem. We implement the algorithm as a sparse matrix vector multiplication on a NVIDIA Tesla K20c GPU. In our implementation we use MATLAB Parallel Processing Toolbox and CUSP library [169], which is a generic CUDA library for sparse linear algebra. Certainly, a more customized and efficient GPU implementation would lead to much faster computation, yet we find ours satisfactory for the 2D and 3D experiments we present. The computation time of each field for a shape of 250.000 voxels is 2 seconds. For 2D shapes, we concatenate sparse matrices of the operators $\Delta - \frac{1}{\rho^2}$ and solve the fields simultaneously, which is not an option for 3D shapes due to memory constraints. The calculation of the projections

for the shapes in 1000 shape database [126] takes approximately 3 seconds.

3.5.2 Boundary Decomposition Based on Regional Information

As described previously in §3.4.1, a natural application of the non-negative sparse decomposition of the shape hyper-field was partitioning of the shape domain into its "meaningful" components. The decomposition is applied to nodes on the shape domain based on their random walk distances to all the boundary nodes as in the demonstration on the hand shape in Fig. 3.8 in §. 3.4.1.

Here, we demonstrate another setting where the shape decomposition is achieved by minimizing the objective function in (3.13) in §.3.4.1 using the transpose of the measurement matrix Y without normalization (3.12). In this setting, in contrast to the previous setting §. 3.4.1, the boundary nodes are decomposed based on their random walk distances to all the nodes in the shape domain. Minimization of the reconstruction error in the objective function depends on boundary nodes and the regions that are associated with each node. That is, the boundary nodes that relate to similar regions are more likely to belong to the same boundary partition. Hence, the resulting decomposition of the boundary inherently depends on a regional partitioning of the shape. An example on a human figure is presented in Fig. 3.11, where decomposed parts and corresponding regions can be observed as active components and basis vectors (respectively) that are factorized from the hyper-field using NNSC.

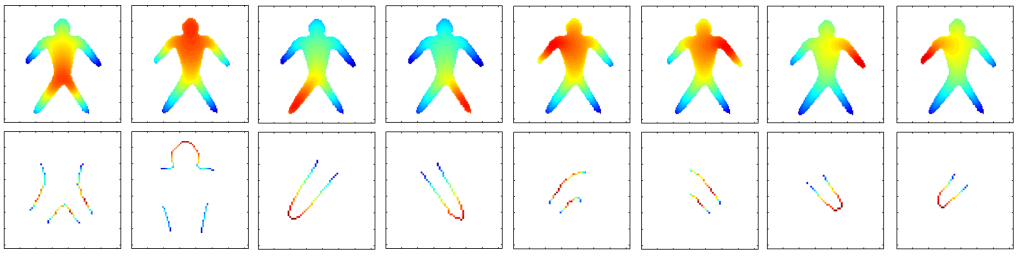


Figure 3.11: Decomposition of the human figure and associated regions. $k = 8$.

Introducing information about regional characteristics of a shape for decomposition of its boundaries leads to rather consistent results. We demonstrate this in Fig. 3.12

for three distinctively different cat poses. The structures after decomposition are very coherent. The sparse decomposition into eight boundary segments reveals the head, the front and rear parts of central body, the tail and four legs. In the third pose only, an additional segment is included in the leg whose regional characteristics are altered due to the significant articulation and deformation, however the inconsistent segment can easily be detected and eliminated considering its low intensity.

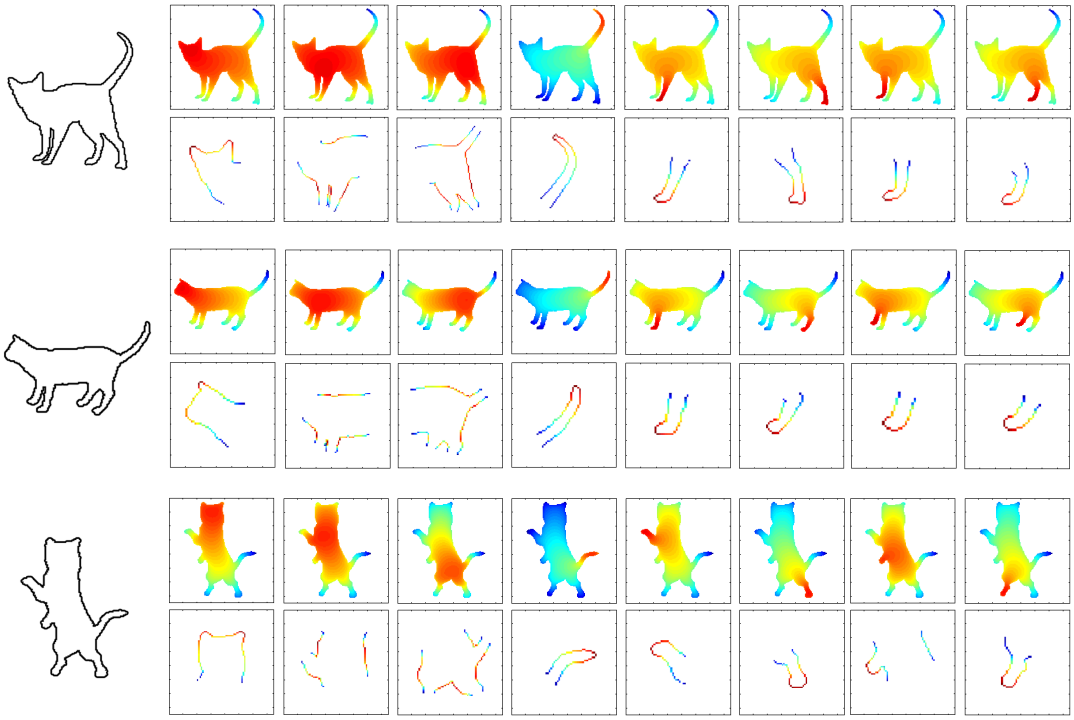


Figure 3.12: Non-negative sparse decomposition over shape hyper-fields of three highly different cat poses partition shape boundary into: the head, the cat frontal body, the back body, its tail, and its legs in a consistent manner.

3.5.3 Orthogonal Projections Based On ρ^2 Sweep: SPEM

For the SPEM, we experimented with a set of shapes that are not necessarily related by isometry. In Fig. 3.13, the projections obtained using the first five principal components are presented for six different cat shapes. The first two projections, which explain most of the variance in the data, are much smoother compared to the remaining projections. In the

first projection, it is observed that the nodes in the vicinity of the boundary attain highly positive values, hence, can be distinguished from the interior nodes, giving only a rough sense of central/peripheral separation. The second projection, on the other hand, exhibits a much stronger central/peripheral separation similar to [1]. The projections(SPEM) obtained using the third or higher eigenvectors encode more subtle details. For example, the third projection reveals ears, head, legs and tail of the respective cat. Notice that these explicitly expressed parts are intuitive and consistent across deformations of the cat shape.

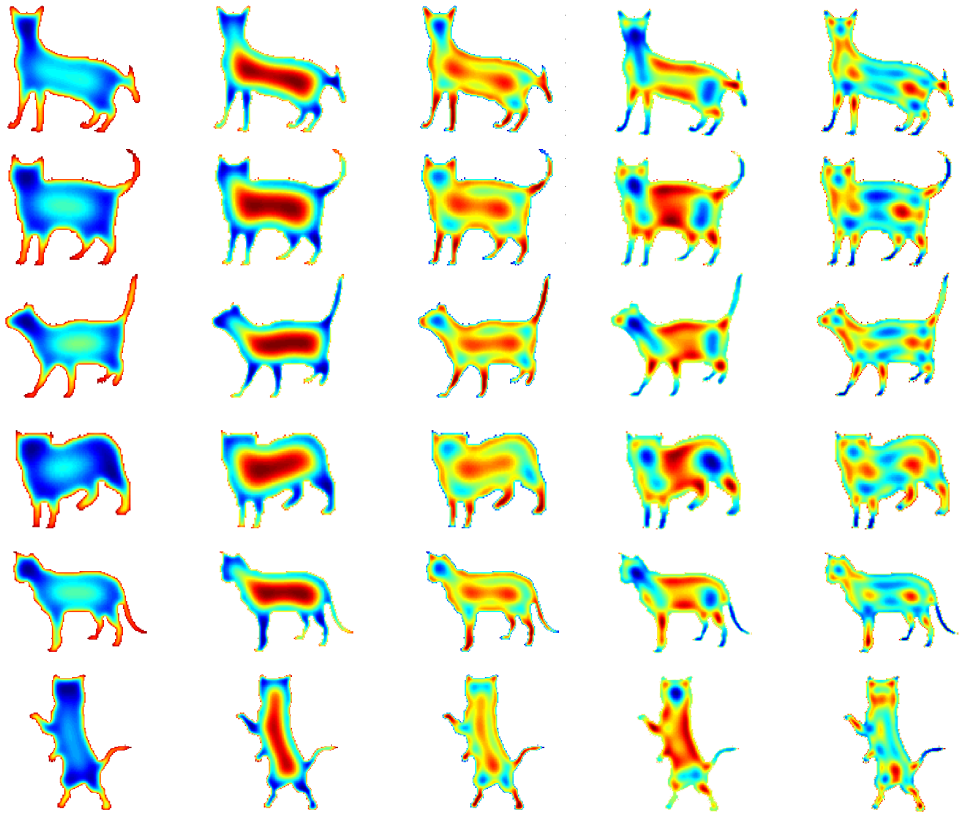


Figure 3.13: Left-Right: First five projections(SPEM): $\mathcal{P}_{1,\ldots,5}$ for 6 different poses of a cat shape, depicted in each row. Each column corresponds to a different projection mode. Hotter colors indicate positive and high values while colder colors indicate negative and low values. Consistency of projections across deformations of the cat shape is observed.

Figure 3.14 demonstrates two things: (i) human figure, with different articulated motion as well as small local deformations shows the projections preserve their character

across those nonrigid deformations; (ii) hand figure with occlusion, local deformation and noise effects show robustness of the projections against noise. The consistency, which can be observed among the projections over each row across the varying instances of the human and the hand shapes, is poised to provide the desired robustness in shape representation required for shape matching and recognition.

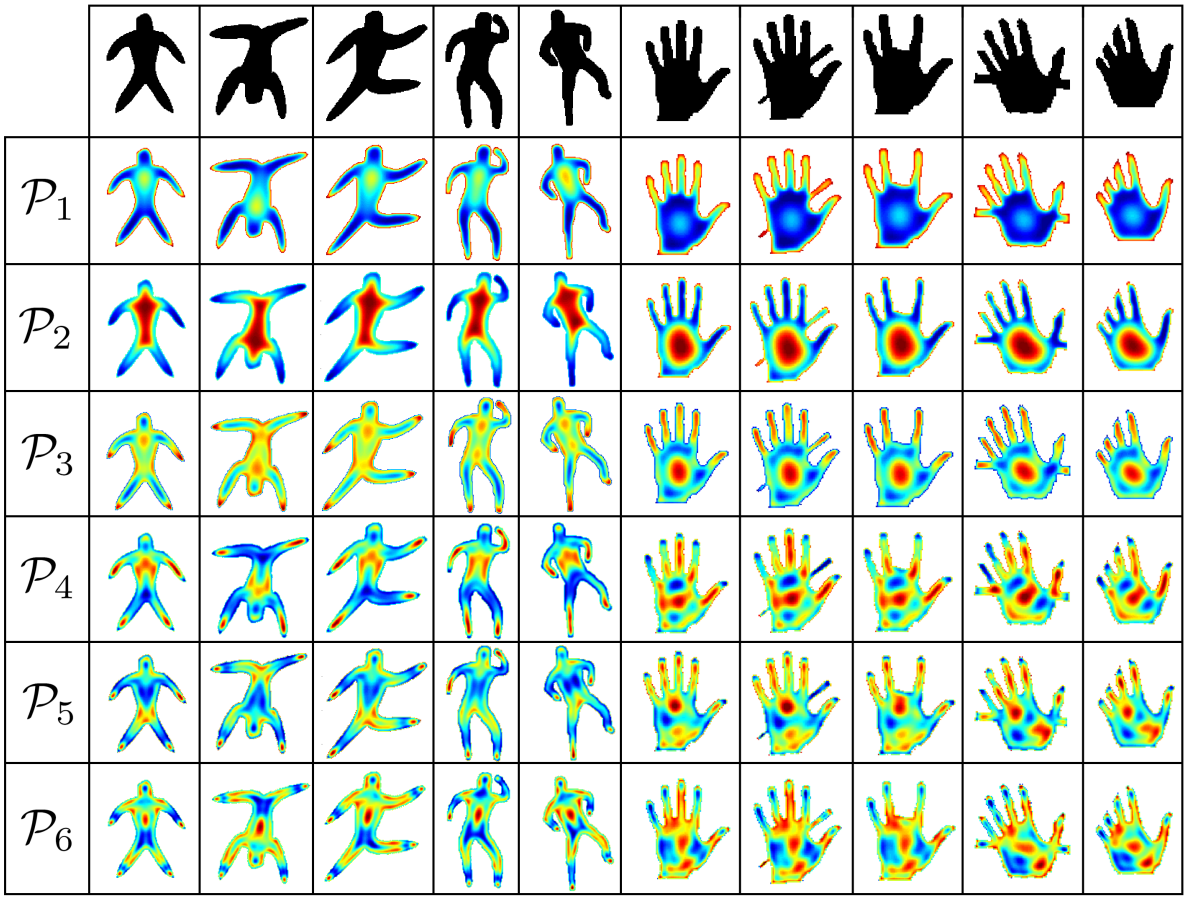


Figure 3.14: First six projections(SPEM): $\mathcal{P}_{1,\ldots,6}$ (on each row) for five different instances (on each column) of a human and a hand silhouette. Human figure displays articulated motion and local deformations. Hand figure displays different noise conditions: occluding a finger; shortening of fingers; protruding two new parts from the hand. Hotter colors indicate positive and high values while colder colors indicate negative and low values. Robustness of projections against occlusion, local deformation, and noise is observed.

In Fig. 3.15, a 3D example of the SPEM is presented. The projections of the 4D hyperfield computed from a 3D horse form onto second to sixth projections are depicted. Since

a 3D form conveys the exact geometry of a real world object as opposed to a 2D shape, which is a perspective projection, our projected fields are naturally more consistent across arbitrary pose changes. In order to be able to visualize different sections distinguished by each projection, we applied a histogram based thresholding procedure. For each projection, one positive and one negative threshold is selected and the surfaces corresponding to the level-sets of those thresholds are visualized. Thresholds were selected at the first jump in the histogram for all the projections and the same threshold was used for the same projections of shapes under different poses. The same remark that was made about the smoothness of the projected fields in the 2D case holds for the 3D case as well. Although some of the thresholded parts can be detached, as in the blue neck part in the sixth projection, the consistency and the similarity of the 3D fields even after the thresholding are notable.

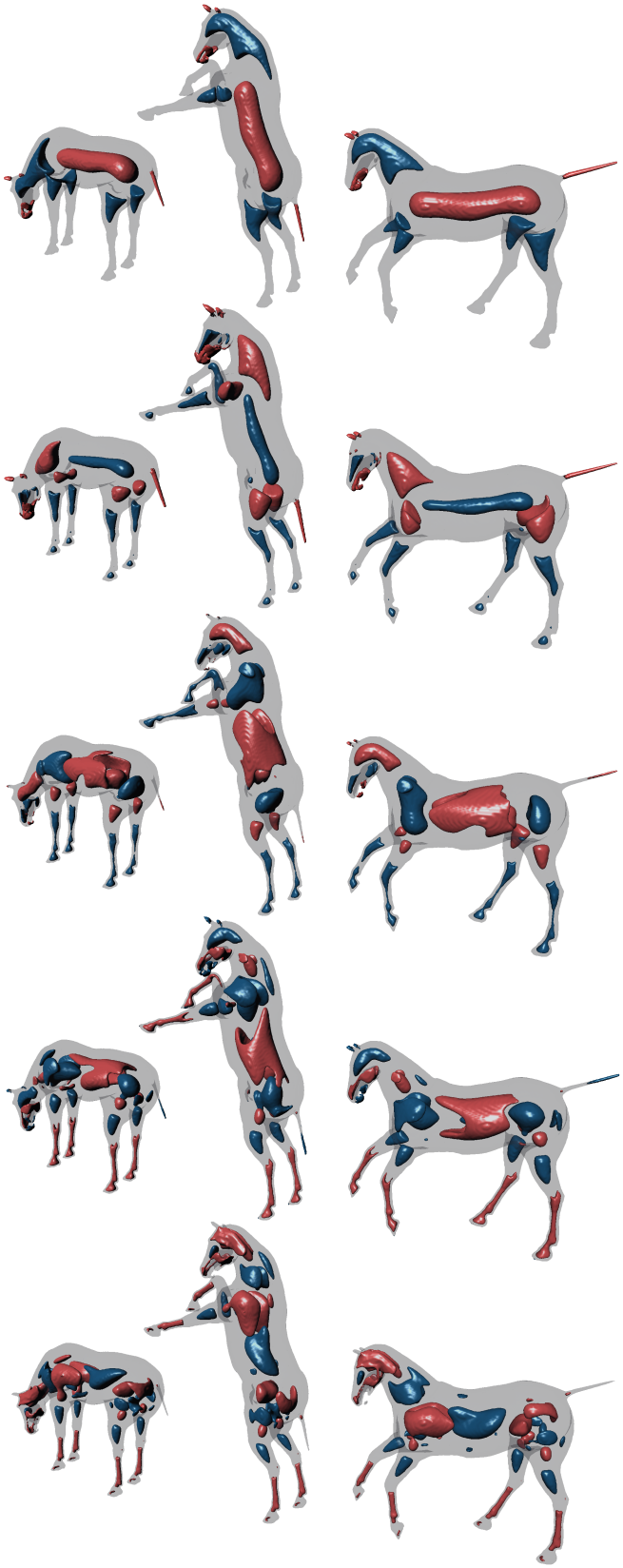


Figure 3.15: Top-Down: Second to sixth projections(SPEM): $\mathcal{P}_{2,\ldots,6}$ for three different poses of a 3D horse. Consistency of each projection across a row for different poses can be observed.

3.5.4 A Moment Based Evaluation of Consistency and Correspondence

In order to quantitatively demonstrate the consistency of the projections, we conducted a classification experiment on the 2D "1000-shape" database, which is an extended version of [126]. The database consists of 50 classes, each containing 20 shapes varying significantly with severe deformations and articulations. From each class, 10 shapes are randomly selected as training data and remaining 10 is used for testing. In order to experiment with the classification performance using a moment-based representation, we extract a group of shapes with disjoint parts from each input single connected binary shape in the database.

The shapes in the database are scaled to a fixed maximum distance to boundary of 20 voxels. Next, the hyper-fields are created, and principal directions are calculated in the ρ dimension. To each projection obtained from a given shape, a basic k-means clustering is applied ($k = 3$) using the intensities of the projections. Thanks to the nature of the SPEM, the resulting cluster centers are very similar: One of the cluster means is very close to zero in terms of projection intensity value and the other two are from the shape nodes that have positive and negative intensity projections, characterizing positive and negative nodal domains. We use the mean of the corresponding cluster for both the negative and positive clusters to generate two new shape maps for each projection. This can be considered as a rough yet straightforward approach for detecting regions that behave similar in ρ^2 space, specifically in a certain principal direction of the hyper-field. We note that a common positive and a negative threshold value is utilized for all shapes in the database. In Fig.3.16, we exemplify the positive and negative shape clusters obtained by thresholding the first five projections on several cat shapes from the database.

As features, we computed Hu's seven invariant moments [170], which are invariant to translation, scale and orientation as features for the classification. The *weak* sense of similarity that these simple moments provide allows us to evaluate the correspondences more clearly. We train linear Support Vector Machines (SVM's) using moments of each generated shape both by stacking features in a cumulative manner and individually to each

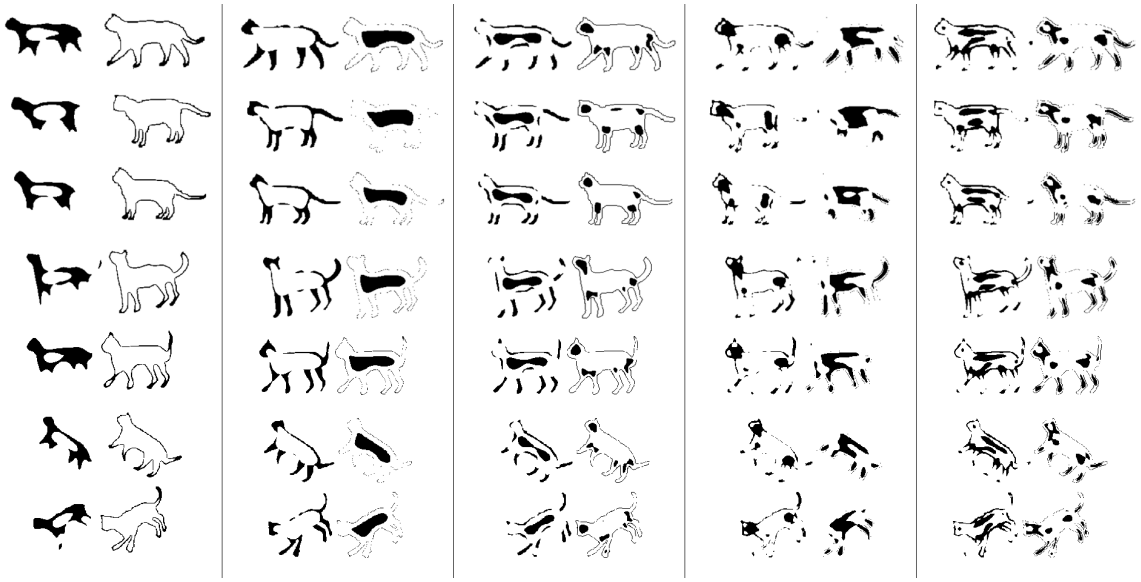


Figure 3.16: Each row contains the negative-positive nodal domain clusters for corresponding to first 5 projections of 7 cat shapes.

projection. The classification results for both experiments that are repeated 100 times, randomized over selection of 10 training and 10 test shapes, can be observed at Fig.3.17.

From the experiments where the SVMs are trained using features from individual shape projections (red), it is clear that the moments for the newly generated shape maps are much more informative compared to only the input shape's moments (blue). This alone shows that level curves of the SPEMs are consistent among shapes of same category and corresponding regions on shapes of the same category have similar projections. The moments obtained from even the ninth projections, which explain very little portion of the variance in the hyper-field, are almost twice as descriptive compared to the original shapes moments (see Fig.3.17). The monotonic behavior in classification performance obtained using the combinations of the moments (black) as features implies that new projections introduce new information (that even the moments can express), which is an observation that is greatly in accordance with orthogonality of the projections.

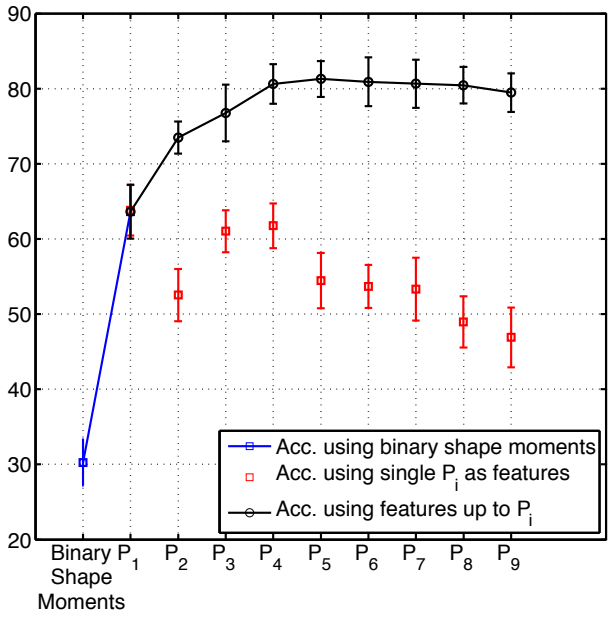


Figure 3.17: SVM Classification accuracies using moment features of: binary shape moments (blue); individual thresholded projections (red); and cumulatively adding thresholded projections (black). Notice that the success rate jumps from %30 (blue) to %80 (blacks) when our approach is used.

3.5.5 Non-Rigid Shape Retrieval Using Screened Poisson Encoding Maps

To address the problem of retrieving similar shapes from a database given a query shape, we utilize the information extracted from the Screened Poisson Encoding Maps (SPEMs). We use a feature-based approach to obtain a compact global shape descriptor from SPEM using feature encoding methods.

An analogy between feature-based 3D shape retrieval and image retrieval was made in [16], where an image is treated as a collection of primitive elements, namely, local image descriptors as *visual words*. The analogy is formed by obtaining *geometric words* using multiscale diffusion heat kernels, which are represented by a geometric vocabulary using soft vector quantization. A similar feature-based approach is used in [146], where Interior Heat Kernel Signatures (iHKS) are used as geometric words with a similar representation

proposed in [16]. Our retrieval approach is mainly similar to those in perspective, yet it differs in the way geometric words are obtained and the way the features are encoded.

As features, we use the SPEMs explained in § 3.4.2. Considering the nature of the problem, due to large variability of the shapes undergoing non-rigid deformations, the features should be robust to bending and articulations, which cause topological changes in the volumetric representation. In Fig. 3.18, we present joint histograms of projections for several shapes that go through large pose changes. The shapes belong to the *SHREC'11* benchmark[87]. The histograms are obtained from the values of SPEMs: $\mathcal{P}_4(\boldsymbol{x})$(horizontal axis) and $\mathcal{P}_3(\boldsymbol{x})$(vertical axis) for all $\boldsymbol{x} \in \Omega$, the logarithm of the number of nodes in the bins are visualized. The choice of fourth and third projections is purely arbitrary, other projections also give coherent results.

The histograms visualized in Fig.3.18 provide only a hint of what the feature space looks like, yet the distinctiveness of the volumetric information encoded is clearly revealed. Even for the shapes that go through large articulated motion and deformation, the representation remains unaltered. Also notice that the representation of woman and man shapes are more similar (yet still distinguishable) in comparison to the representation of other shapes that are less related.

In order to compactly represent a shape for the retrieval application, we use the feature encoding method: Vector of Locally Aggregated Descriptors (VLAD)[88]. VLAD characterizes the distribution of vectors with respect to the pre-computed centers, words that belong to a vocabulary. Unlike the *Bag of Features* or *soft vector quantization* approaches, where the distances of the features to centers are accumulated, the difference *vectors* from each feature to assigned center are aggregated.

For a node in the shape domain $\boldsymbol{x} \in \Omega$, the SPEM $\mathcal{S}(\boldsymbol{x})$ consists of the projections $\mathcal{P}(\boldsymbol{x})$: $\mathcal{S}(\boldsymbol{x}) = \{\mathcal{P}_1(\boldsymbol{x}), \ldots, \mathcal{P}_d(\boldsymbol{x})\}$ where d is the number of projections used. A codebook, $\mathcal{C} = \{\boldsymbol{c}_1, , \ldots, \boldsymbol{c}_k\}$, of k representative points $\boldsymbol{c}_i \in \mathbb{R}^d$ is acquired using k-means clustering algorithm. Each node in the shape domain is assigned to the nearest cluster center, as a hard vector quantization to obtain $\mathcal{NN} \in \mathbb{R}^d$:

$$\mathcal{NN}_{(\boldsymbol{x})} = \underset{\boldsymbol{c}_i \in \mathcal{C}}{\arg\min} \|\mathcal{S}(\boldsymbol{x}) - \boldsymbol{c}_i\| \tag{3.15}$$

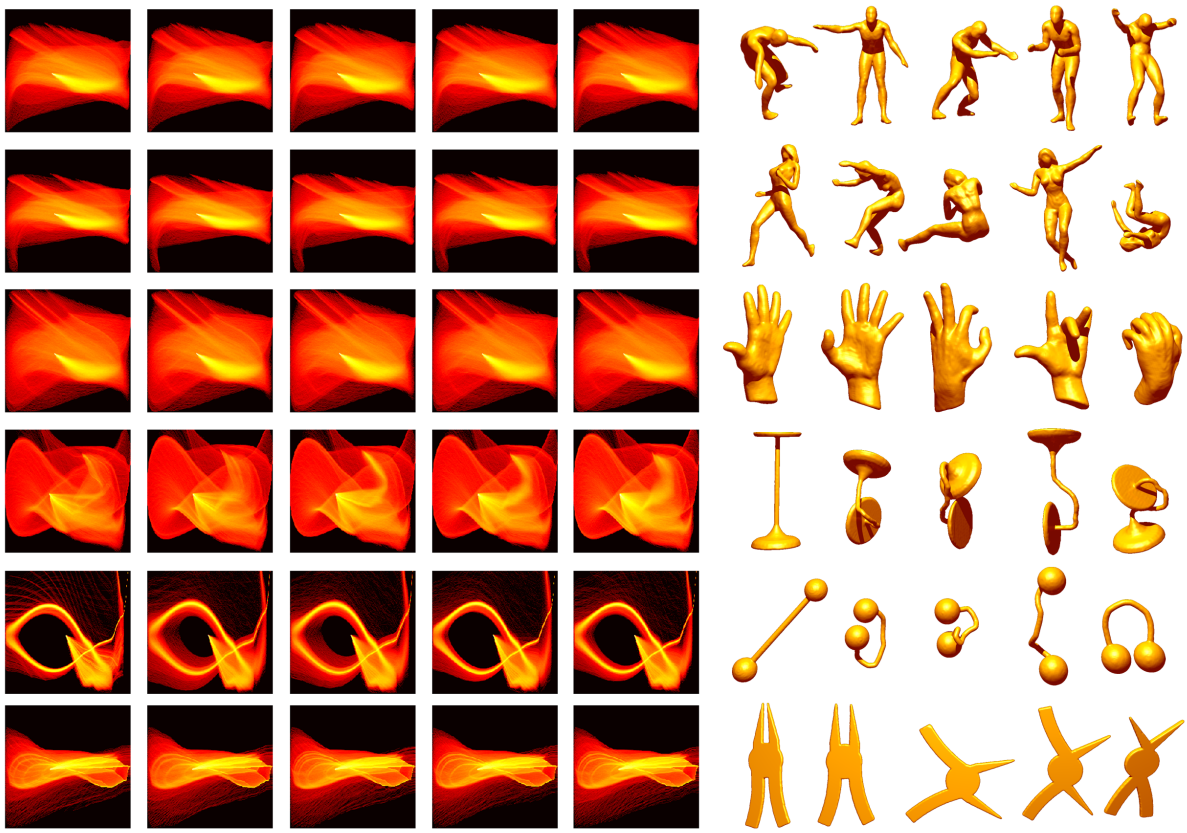


Figure 3.18: Joint histograms inside SPEMs: $\mathcal{P}_4$ vs $\mathcal{P}_3$ for corresponding shapes on the right. The histogram intensities are displayed using a logarithmic scale. The articulations have almost no effect on the joint histograms and there is large variation in histograms of shapes with different volumetric structures.

The differences $\mathcal{S}(\boldsymbol{x}) - \boldsymbol{c}_i$ of the vectors assigned to each center $\boldsymbol{c}_i$ are accumulated to obtain d dimensional *residual* sum vectors $\boldsymbol{R}_i$:

$$\boldsymbol{R}_i = \sum_{\boldsymbol{x} \in \Omega \colon \mathcal{NN}(\boldsymbol{x})=\boldsymbol{c}_i} \mathcal{S}(\boldsymbol{x}) - \boldsymbol{c}_i \tag{3.16}$$

The aggregated residual vectors $\boldsymbol{R}_i$ are then normalized so that they have unit L_2 norm. This is followed by a power normalization, which can be considered as a variance stabilizing transform:

$$\theta_i = sign(\boldsymbol{R}_i)\left(\frac{\boldsymbol{R}_i}{\|\boldsymbol{R}_i\|}\right)^{\alpha} \tag{3.17}$$

where the power operates element wise on the given vector. This normalization helps cop-

ing with the undesired burstiness of the encoded vectors. One of the sources of burstiness is the different number of nodes assigned to each center. In [171], several approximations are introduced that associate (3.17) to earlier work on variance stabilizing transforms on a compound Poisson distribution [172, 173]. Another source of variance in our case is the differences in the intensities of the projections. The SPEM projections that correspond to eigenvalues that express little variance have much lower intensities, hence little contribution in comparison to, say, $\mathcal{P}_1$. The power transform enhances their contribution in the overall representation. The vectors θ are concatenated and again normalized to have unit L_2 norm as a vector of length $k \times d$ to form the final representation of a shape.

We conducted the retrieval experiment using the *SHREC'11 Shape Retrieval on Non-Rigid 3D Watertight Meshes* benchmark[87]. The database consists of 30 classes, each with 20 samples that inherit large intra-class variation in terms of articulated motion and non-rigid deformation. In our implementation, all of the shapes in the database were normalised to a constant scale of $\sim 250,000$ voxels. We used first six projections $d = 6$, and trained a vocabulary C of size $k = 32$ words, resulting into a 192 dimensional vector as a global descriptor of shape. The vocabulary is learned from 50000 randomly sampled nodes from each shape in the database. Standard deviations for the SPEM projections that belong to each projection index are also estimated using this training data. While evaluating, each SPEM projection is first L_2 normalized then multiplied by the standard deviation estimated for that projection index. The nodes that are next to the boundary are discarded while encoding, since they are less informative and might get affected from boundary discretization. The idea is to normalize the projection intensities, without changing the intensity range of the projections. We observed that such a normalization causes an increase in retrieval performance. For power normalization of VLAD, we use $\alpha = 0.25$ in Eq.(3.17). Correlation distance between the 192 dimensional global descriptor is used as a measure of dissimilarity between query shape and all shapes in the database.

The retrieval performance is demonstrated in Table 3.1 using five standard retrieval statistics: nearest neighbor (NN), first tier (FT), second tier (ST), E-measure (E), and discounted cumulative gain (DCG). For the details regarding the measures, we refer to [174]. Results for best four methods from the participants in the contest are presented.

The results for these are taken from [175], where the methods are also briefly presented. Additional results in the table are taken from the work of Rustamov: iHKS [146] and Konukoglu's spectral distance method: WESD [58]. Hybrid method results are obtained the way *SD-GDM + MeshSift* hybrid method is presented in [87], specifically, by applying a min-max algorithm to distances obtained using two methods and adding them up to obtain the final distance matrix. We merge our results with other methods in the proposed manner, using the distance matrices participated in SHREC contest, to demonstrate how combination of our work and existing work performs. We also present the precision-recall performance of our retrieval approach and the top four contestants in SHREC'11 track in Fig.4.5.

Table 3.1: SHREC'11 Retrieval on Non-Rigid 3D Watertight Meshes Database Results

		NN	**FT**	**ST**	**E**	**DCG**
Methods:	Our Method (SPEM)	99.8	**97.4**	**98.6**	**73.3**	99.3
	SD-GDM [176]	**100.0**	96.2	98.4	73.1	**99.4**
	iHKS[146]	99.5	92.2	95.5	71.0	98.0
	MDS-CM-BOF[177]	99.5	91.3	96.9	71.7	98.2
	ShapeDNA[59]	99.2	91.5	95.7	70.5	97.8
	WESD[58]	99.3	90.2	93.05	69.0	97.1
	MeshSIFT [178]	99.5	88.4	96.2	70.8	98.0
Hybrid	SPEM+SD-GDM	**100.0**	**98.4**	**99.5**	**74.0**	**99.7**
Methods:	SPEM+MDS-CM-BOF	100.0	97.5	99.1	73.7	99.6
	SPEM+MeshSift	99.7	97.6	98.9	73.5	99.5
	SD-GDM+MeshSift	100.0	97.2	99.0	73.6	99.6

Our method outperforms existing methods in terms of first-tier, second-tier and E-measure performance. It ranks the second for the NN and DCG measures with very close numbers. Only one of the nearest neighbor results out of 600 is retrieved falsely, which

is a sample from the *ants* class. Generally, the method only fails when internal distance relations undergo large changes, which occasionally happen in ants class deformations. Among the methods compared, only iHKS uses volumetric information. Our method clearly outperforms iHKS with respect to all of the performance measures.

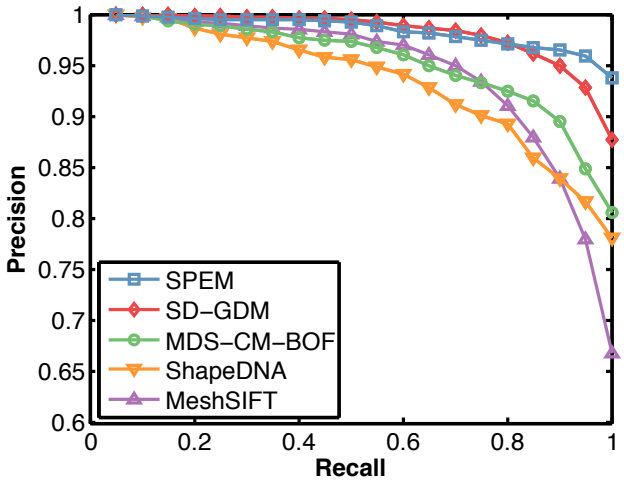


Figure 3.19: Precision - Recall Performances in Shrec'11 Non-Rigid Database

The results suggest that extracting volumetric information in a robust way can lead to enhancement in the non-rigid shape retrieval performance when compared to extracting information regarding only intrinsic surface properties. Moreover, as expected, combining volumetric information and surface information results in a significant boost in performance, as observed in resulting hybrid performance of SPEM in Table 3.1.

3.6 Conclusion

In this paper, we provided a novel distance hyper-field representation for shapes via screened Poisson PDE. An acknowledged advantage of the presented shape representation, which is valid for all distance-based shape representations, is that the shape model is free from the dimensions of the shape: the volumetric description of the hyper-fields we propose is readily extendable to any shape embedding dimensions in $\mathbb{R}^n$ for $n = 2, 3, 4....$

For an $n - d$ shape with m boundary points, the hyper-field is a scale-space stack

parameterized by the screening parameter, where each member of the stack v^ρ is a superposition of m fields, each of which is the solution of the PDE with a point source placed at a boundary point. The shape hyper-field captures all sorts of characteristic information within and on the boundary of the shape, therefore encodes both local-global and interior-boundary interactions.

The new representation is also tied to a collection of conditioned random walks, each walk emanating from a fixed boundary point and *walking* with a bias controlled by the screening parameter; and, the effect of the screening parameter is a change of measure.

Furthermore, we demonstrated extraction of shape information from the hyper-field, which is a rich high-dimensional representation, by two felicitous compact decompositions to exemplify: (i) natural shape partitions (NNSC); (ii) consistent shape maps through SPEM. The potential of extracting various shape descriptors from the introduced shape hyper-field was demonstrated over both a 2D "1000-shape" database[126], and a benchmark dataset *SHREC'11* [87]. The SPEM performance was evaluated by using the VLAD method [88] for volumetric feature encoding. The SPEM consistently ranked the first or the second in all measures, and ranked the first when a hybrid combination with top surface-based methods was computed. Another interesting property of the SPEM was its adaptation to scale, which was experimentally verified by its performance over the benchmark.

The shape hyper-field representation presented in this paper is constructed over the whole shape domain, therefore, it is certainly possible to either define new shape measures or adopt existing popular descriptors, even image descriptors, in order to apply the introduced hyper-field framework to shape matching, classification or partitioning problems.

4

Shape Matching using Image Descriptors

4.1 Introduction

In this chapter, we will introduce the SIFT-based image matching framework, which is for the first time adopted to the planar shape-matching problem. This proves to be an intuitive and effective method for which performance will be demonstrated compared to state-of-the-art shape retrieval methods over various shape databases.

4.1.1 Related Works

The idea of matching shapes from the same class using correspondences that can be related via geometric transformations, stemming from the work of D'Arcy Thompson [179], motivates landmark-based methods for shape matching. Typically, the landmark points are acquired by sampling either from shape boundary [180, 55] or medial axes [181, 182, 85, 46, 183, 50]. Such landmark-based shape representations have been succesfully used in shape matching, also several robust methods for point-based matching are presented [184, 185, 186, 187].

In a different and very active line of research, local detectors that are covariant to a class of transformations as support regions to compute invariant descriptors have proven

to be very successful in image retrieval and object recognition [188, 189, 190, 191, 192, 193]. Matching using such a representation is commonly followed by a spatial verification procedure [194], where a planar homography transformation hypothesis is formed and agreeing matches are kept as inliers. This powerful framework progressively evolved over the years and adopted for matching images of objects.

There are, nonetheless, applications where matching silhouette data rather than image data is needed. For example, in computational anatomy, alignment of shapes (via silhouettes or silhouette boundaries) is crucial in constructing anatomical atlases for organs or characterizing change that may be a precursor to certain diseases or defects. Furthermore, availability of depth images via RGB-D cameras made it possible to extract the silhouette data for problems where color and texture may be uninformative. Hence, adopting the highly evolved body of works available for matching images to matching silhouettes is important for not re-inventing equivalent tools for silhouette matching.

4.1.2 Our Contribution

The reason that this framework has never been proposed to work with silhouettes of objects is the lack of a shape representation that encodes pure shape information on the shape domain and is both *consistent* and *expressive*. On one hand, smooth distance fields given for instance by solutions to the Poisson equation are consistent, but they are highly smooth and not sufficiently expressive; hence, the schemes based on Poisson type shape representations employ secondary level representations in the form of skeletons [126] or descriptors such as weighted moments [52]. On the other hand, representations based on Laplace eigenfunctions are richer yet inconsistent for non-isometric changes unless jointly computed for pair of shapes to be matched [144].

We propose that the SPEM features in the form of a shape field are both consistent and expressive so that descriptors and geometric assumptions proposed for object images can be used for object silhouettes. Via the presented field representation utilizing the whole shape domain, matching becomes possible without being plagued by the high-dimensionality and redundancy of region-based representations. We demonstrate the effectiveness of our representation using SIFT descriptors by performing shape retrieval

experiments on widely experimented common shape datasets and compare retrieval accuracy using precision/recall and Bull's eye scores. In all the experiments our scheme outperforms the related work.

The solution to the screened Poisson equation, v, is an exponentially decaying smooth distance function which has been shown to implicitly code curvature in [85], where the first time the model has been proposed as a means of shape representation. Both [85] and the follow-up work utilised v for extracting more abstract representations such as shape skeletons or parts. None of the works approached the problem as densely sampling the parameter space to obtain consistent and expressive maps(SPEM), hence, translating the shape matching task into image matching task for which a highly evolved rich machinery is available.

4.2 2D Shape Matching and Retrieval Using Consistent Projections

Due to the orthogonality aspect of the PCA, the values of the projections are fluctuating between positive and negative. The nodal domains emphasize *internal parts* of shapes (not necessarily semantic). Within these parts, the behavior of a shape point under varying ρ are similar in the direction of a specific principal component. These internal parts can be analyzed as blobby structures in $\mathcal{P}_i$. A local extemum of a projection can be considered as an approximate center of an internal part, where the specific behavior in ρ space is the most prominent. We propose a shape matching procedure where corresponding internal parts of shapes are matched under the constraints of having *similar projection characteristics* and a *geometrically likely configuration.*

The first step toward shape matching using projections is detecting regions that are relevant (local feature frames). In our application, we used the Hessian-Laplace detector along with the affine adaptation process [190][191]. Many scale and affine invariant region detectors have been proposed and a comparison can be found in [191]. Among these, the ones of interest to us are those designed for *blob* detection: In [195], Lindeberg used the maximum of the normalized Laplacian in scale-space, whereas Lowe [89] approximated the

Laplacian with difference-of-Gaussian(DoG) to detect local extrema in scalespace. Lindeberg and Garding [196] also introduced an affine adaptation process that makes use of the second moment matrix. Hessian-Laplace region detector [190] is invariant to orientation and scale changes. The local maxima of the determinant of the Hessian indicate a presence of a blob structure. Scale selection is via Laplacian-of-Gaussian. Hessian-Laplace scheme yields higher localization accuracy in scale-space compared to DoG detector. Moreover, the number of detections returned by the Hessian-Laplace detector is larger, leading to enhanced information for geometrical assessment. Overall, our choice of the Hessian-Laplace as the detector is supported experimentally. Considering that the forms of the internal parts defined by projections can be better estimated as ellipses rather then circles, further justifies the choice of affine adaptation. In order to detect signal variations that lie on the blob boundaries, the measurement region is set as three times larger than the detected region. The projections are zero mean and the regions outside shape boundaries are taken as zeros. The detected regions may also be centered outside the shape domain point. This might at first seem counter-intuitive; however, the exterior regions close to the shape boundary also encode shape.

Detection of local feature frames is followed by generating descriptors that are well suited for the matching process. We employ the SIFT descriptor of dimension 128 in its original form with Hessian-Affine support regions. This is calculated using VLFeat library [197]. In the matching process, an initial set of correspondences are established using the 2NN (2-Nearest-Neighbor) [188], and then refined using a RANSAC based [90] scheme under planar homography assumption. The idea is to retain the correspondences that are in accordance with the estimated homography as inliers while discarding the rest. Finally, inliers detected for each projection pair are combined to yield the final refined set of correspondences. In Fig. 4.1, 2NN matches between two cat shapes are presented. Fig. 4.1 (a) depicts the initial set of correspondences, while (b) and (c) depict final set of correspondences refined via relaxed and strict homography assumptions, respectively. The total number of retained correspondences is employed as a similarity measure. At a first glance, it may seem that a planar homography assumption would not be sufficient to represent deformations that a shape may go through, especially articulated motion. Nev-

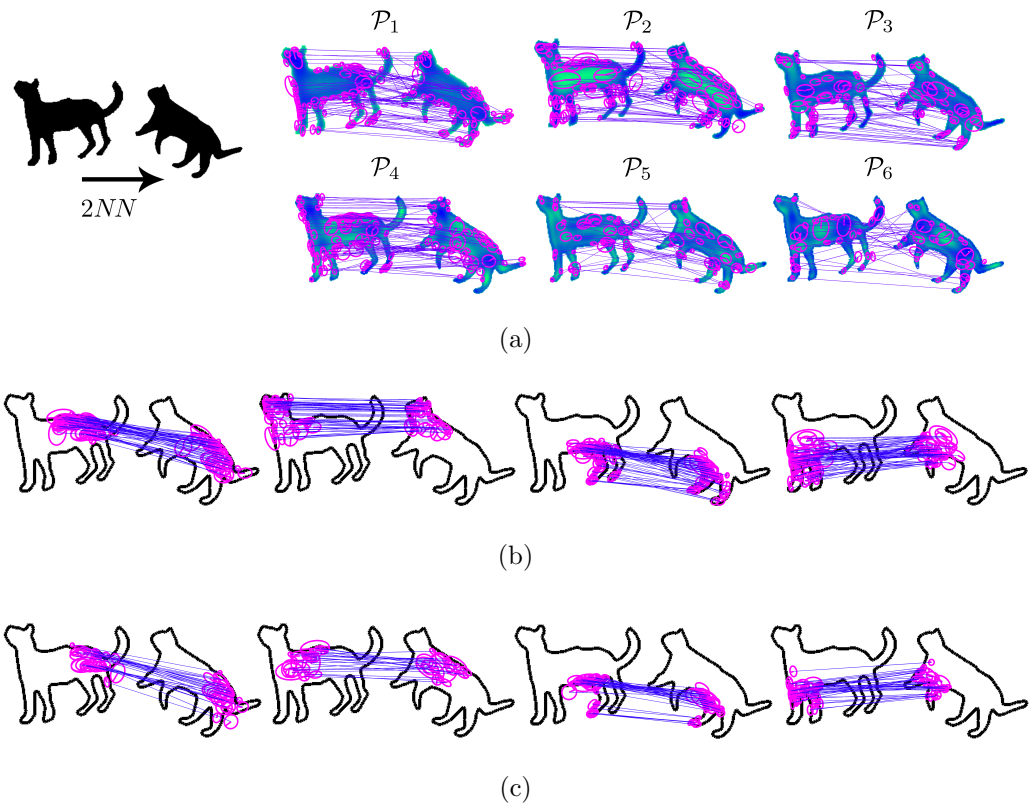


Figure 4.1: (a) 2NN SIFT Matching results for $\mathcal{P}_i, i = 1, 2, .., 6$ for two cat shapes. (b,c) Refined matches after geometric verification with a (b)*relaxed* (c) *strict* planar homography assumption. For solely visualisation purposes, the matches are clustered using a k-means algorithm based on spatial distances.

ertheless, the existence of the articulated part is encoded in the representation, not only on the shape locations belonging to articulated parts but also on central parts. Furthermore, the comparison between two local feature frames of which centers lie in an internal region also uses the information from possible articulated structures nearby since the feature measurements are from a much larger area than the detected region. The ellipse-like (elongated) shape of the blob contributes to the robustness, too. For the evaluation of similarity, the relaxed homography assumption acts rather as a regularizer. Especially considering that planar shapes are perspective projections of non-planar objects, the homography assumption is well suited. The key to good performance under a broad range of geometric configurations is the consistency of the underlying representation. Notice that there are large number of matches where the limbs connect to the main body and

near the neck even for the strict homography assumption in Fig.4.1 (b). It is less relevant whether we can match the articulating head or not, as long as we can match the neck region which contains information that diffuse from the head and as long as SIFT detectors notice that there is the head outside the detected region which has gone through an affine transformation.

In our scheme, unlike in several other methods such as SC [180], IDSC [55], graph editing [50], how much a shape is transformed does not affect the similarity score (other than refining matches). That is we do not penalize deformations due to perspective projections. This smoothly extrapolates to wider range of visual transformations: In Fig. 4.2, four illustrative matching examples are depicted. The first pair is two jet planes that cannot be related by a group transformation. The local structures are very different and the matching is further complicated by an occluding part in the input shape. In the bottom two, the input shape (shown in green colour on the left) is projected on the target (shown in blue colour on the right). The projection of the green target using the estimated homography is shown as red colour drawn on the target. We remark that the regions that are locally different due to occlusions play no role in the matching procedure and do not contribute to the similarity measure. The second pair is two wrenches, one with a single open end and the other with two. In this case there are two alternatives, both of which will yield a partial match. The one which yields relatively a better match is chosen. In the third example, a pair of shark shapes are given. Majority of the matches are due to rear sections, which contain the most partially similar structures. The method's ability to handle partial matches is also visible in the last example. Notice that almost no information is employed from the tail regions where the two shapes are dissimilar. These illustrative examples provide an intuition on how the homography assumption corresponds to a regularization that favours the best possible partial match.

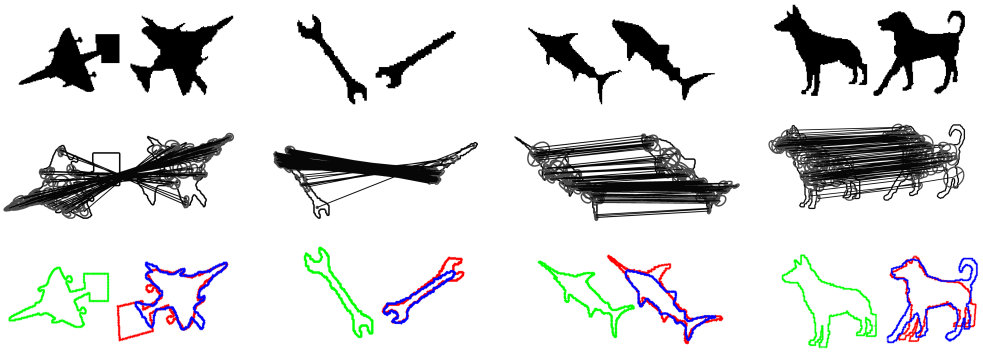


Figure 4.2: Matching results for four pair of shapes. *First Row:* Input and target shapes. *Second Row:* Inlier matches within planar homography. *Third Row:* Input shape projected onto the target using the estimated homography.

4.2.1 Shape Retrieval Approach

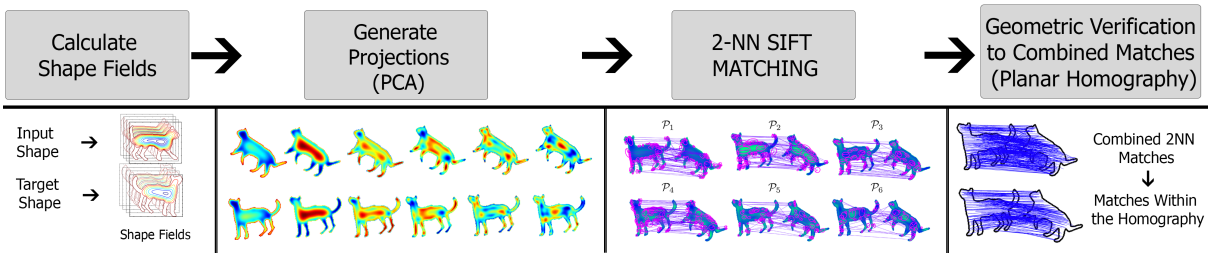


Figure 4.3: Proposed Shape Matching Scheme

In order to find a pairwise similarity measure between a given query shape and all of the shapes in a given database, we utilize the matching procedure depicted in Fig.4.3. All shapes are scaled to have the same maximum distance to boundary (maximum distance transform) prior to the generation of the projections. In the course of our experiments, we have observed that the strictness of the homography assumption (allowed deviation from hypothesis to be accepted as an inlier) does not significantly affect the overall retrieval accuracy as long as the same assumption is used in all geometric verifications. For the 2NN test [89], we declare a match if the distance ratio between first and second closest neighbors are below 0.85.

The individual contributions of the projections in the matching procedure involving the first nine projections are presented separately for shapes of different categories(left) and shapes of the same category(right) in Fig.4.4. The contributions are depicted as the

distributions of the ratio of the 2NN matches of a specific projection among all matches (x-axis) and ratio of matches that are in agreement with the geometric verification hypothesis among 2NN matches (y-axis). The distributions obtained from the 180-shape dataset [126] are approximated with Gaussian functions and displayed with standard deviation $\sigma = 1$. Notice that the ratio of inliers is significantly larger for the matches between the shapes of same category. Also note that the contribution of each projection decreases as the projection index increases. When both the match contributions ($\frac{2NN(\mathcal{P}_i)}{2NN_{Total}}$), and inlier ratios ($\frac{INLIERS(\mathcal{P}_i)}{2NN_{Total}}$) are relatively large, they would indicate that the informativeness of a projection is high. Larger indexed projections, e.g. $\mathcal{P}_9$, have lower inlier ratios, which have similar characteristics regardless of whether the shape is from the same category or not. Notice the overlap of $\mathcal{P}_9$ samples on the right to those on the left. In the experiments, we use projections $\mathcal{P}_1, \ldots, \mathcal{P}_7$ for the matching procedure. This choice is verified experimentally as the retrieval accuracies stabilize around $\mathcal{P}_7$, and start to drop slightly after $\mathcal{P}_9$, which is in line with the observations in Fig.4.4.

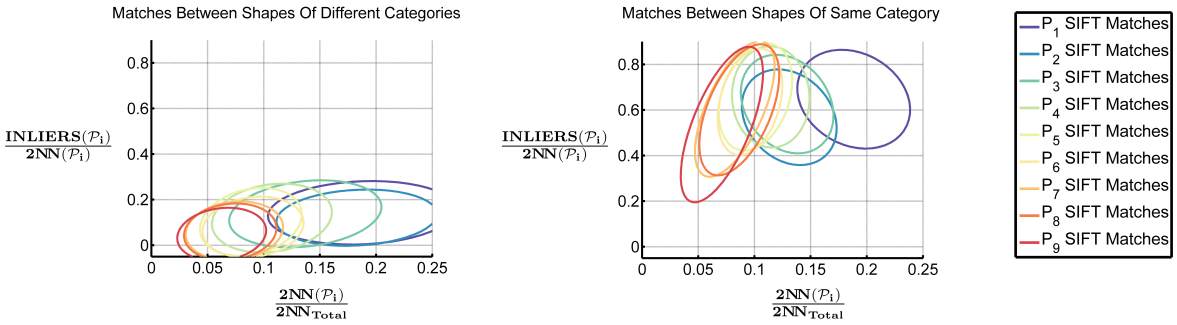


Figure 4.4: Distributions of match contributions versus inlier ratios from each projection. Left: Matches between shapes of different categories. Right: Matches between shapes of same category.

4.2.2 Shape Retrieval Experiments

In this section, we present our results in comparison to results of the state-of-the-art shape retrieval methods on three commonly employed datasets of sizes $490, 180$ and 1000. The selected methods are Poisson features, MPEG-7 descriptors, as well as recent axial or boundary landmark-based representations. We also experiment with the methods

based on a modern idea of going beyond pairwise similarities. These methods learn the manifold structure of shapes in the database by further processing of the similarity matrix, in particular, by spreading the affinities on the set of shapes. There are several successful methods for imposing beyond pairwise information to retrieval task e.g. [198, 199, 200, 201, 202, 203]. We will refer to the collection of methods that employ more than a pair of shapes in defining pairwise similarly as context based similarity methods and will include several variants in this group.

As reported in Ling et. al [204], improving pairwise similarity measurements does not always imply a better bull's eye score. *Perceptual* closeness of the retrieved shape to the query is also critical to form a reasonable graph structure constructed by similar shapes to be further processed by affinity spreading in order to learn the manifold structure of shapes in the dataset. To point out that our pairwise approach provides successful perceptual resemblance to the query, we also report our results for the context-imposed retrieval by using the method of Donoser and Bishof [203].

4.2.2.1 490-Shape Dataset

The 490-Shape Dataset [52] is a collection of natural silhouettes expanded by variable classes from Kimia Dataset [205]. It contains 12 classes with unevenly distributed samples. In [52], features that are extracted from a Poisson field are employed to compute pairwise shape similarity for matching. These Poisson features are called PF for short, and they are weighted inversely proportional to the ratio of the within class variation and between class variation. Then the weighted Euclidean distances between features are used as a measure of shape similarity. In [52], the performance is measured as the number of correct matches in the top 1 and 15 retrievals. Based on this performance measure, it has been compared to geometric moment descriptors (GMD) and Zernike moments descriptors (ZMD) using weighted Euclidean distances and city block distances [206]. It has also been compared to the MPEG-7 descriptors: Angular Radial Transform Descriptor (ARTD) [207, 208] and Curvature Scale Space Descriptor(CSSD) [209][208].

To be comparable, we repeat the same experiment and report comparative results in Table.4.1. We obtain nearly %15 improvement over Poisson features (from %84 to %96.05).

Table 4.1: Retrieval Performance for 490 shape database for 15 closest shapes.

	ARTD[207]	GMD[206]	CSSD [208].	ZMD [206]	PF[52]	Our Method
# hits in top 1	86	89	89	93	93	**99.59**
# hits in top 15	64	72	76	77	84	**96.05**

Our method also significantly over performs Poisson features in terms of precision-recall values. We can only compare up to nearly %30 recall since the results beyond this recall level are not reported. At %30, our method using SIFT features yields %96 percent precision whereas the method based on Poisson features yields %85. At %20 recall, the precision with our method reaches to nearly %100 whereas the precision of the Poisson features remains around %87.

As a final performance measure, we plot complete precision-recall graph for our method using SIFT features (Fig. 4.5 (left)). At high recall levels, the precision is lower compared to other datasets. This is because the number of shapes in most confused classes (cats, dogs and horses) are larger compared to the number of shapes in the other classes. This does not pose a problem if the same number of shapes are being retrieved for all queries (as in Table.4.1), but lowers the overall precision at high recall levels. To examine further, we perform a second retrieval experiment where all the four-legged animals are combined and considered as a single class as done by Kimia [205]. After this lumping, the retrieval performance increased drastically. If we further impose context using [202], the precision at %100 recall reaches to %97.40 using original classes and %99.47 using combined classes.

4.2.2.2 180-Shape Dataset

This dataset by Aslan and Tari [126] includes 6 shapes from 30 categories. Within each category, the samples go through significant articulations in addition to changes in scale and orientation. This is the most articulated shape dataset, thus, a planar homography transformation assumption is expected to fail. Nonetheless, as we have argued in Section 4.2 introduction, due to the high expressive power of our representation and descriptiveness of SIFT, the partial correspondences accepted within homography are quite capable of distinguishing shapes of different categories.

The retrieval performance of our method is compared to those of Inner Distance Shape Context (IDSC) [55], Disconnected Skeletons [126] and its context induced version [210], as the highest performance on this dataset has been reported by these three methods. Our method outperforms all the three (Table.4.2). At %100 recall, our precision is %97.69 and non of the top 3 retrievals contains error, while the precision of IDSC [55] is %95.83. With the context information, the precision reaches to %100 at full recall.

4.2.2.3 1000-Shape Dataset

The 1000-Shape dataset is an extended version of the Aslan-Tari dataset. It consists of 50 classes, each containing 20 shapes with large deformations and articulations. The retrieval performance is again compared to the state-of-the-art boundary and axis based methods (Table.4.3, Table.4.4, and Fig. 4.5 the middle and the right). The scores for Shape Context(SC) [180] and Inner-Distance Shape Context (IDSC) [55] where the affinities are computed using dynamical programming [55] are taken from [211]. The scores for Aspect Shape Context(ASC) descriptors are from Ling et. al [204] who presented the best available result on Tari-1000 dataset by using the context information. As reported in Table.4.3, when only pairwise matches are considered, our method yields the highest precisions at all the four recall levels, though the performance of shape context based descriptors are still comparable. In Table.4.4, we report scores for ASC improved with Locally Constrained Diffusion Process(LCDP) [199], SC and IDSC improved using Label Propagation(LP) [201] and Tree Edit improved using category knowledge [210]. For each method, the precision at 100% recall as well as the Bull's Eye score are listed. The complete precision-recall graphs are given Fig. 4.5 for both pairwise and context imposed experiments.

Table 4.2: Retrieval Performance for 180-shape dataset

	Tree Edit[210]	Disconnected Skeleton[126]	IDSC[55]	Our Method
at 100% recall	87	88	95.83	**97.69**

Tree Edit +Context[210]	Our Method +Context
99.41	**100**

Table 4.3: Retrieval Performance for 1000-shape dataset

	Tree Edit[210]	SC[180]	ASC[204]	IDSC[55]	Our Method
at 25% recall	88	98	98	97	**98.62**
at 50% recall	78	95	96	94	**96.61**
at 75% recall	65	91	90	91	**93.45**
at 100% recall	31	85	68	84	**87.45**

Table 4.4: Retrieval Performance using context methods for the 1000-shape database.

	Tree Edit +Context[210]	SC+LP [180] [201]	IDSC +LP [55] [201]	ASC +LCDP[204]	Our Method +Context
at 100% recall	51	93	94	95	**98.18**
$Bull'sEye$	-	97.79	99.35	99.79	**99.91**

4.2.3 Discussion

To summarize, evaluations of the proposed shape matching method were performed on three common datasets. Comparisons were against methods that used regional features such as Poisson features [52] or more compact landmark-based methods such as [180, 126] to exemplify two main categories of shape descriptors. For the 490-shapes dataset, our method improves the retrieval performances with respect to those of the closest Poisson

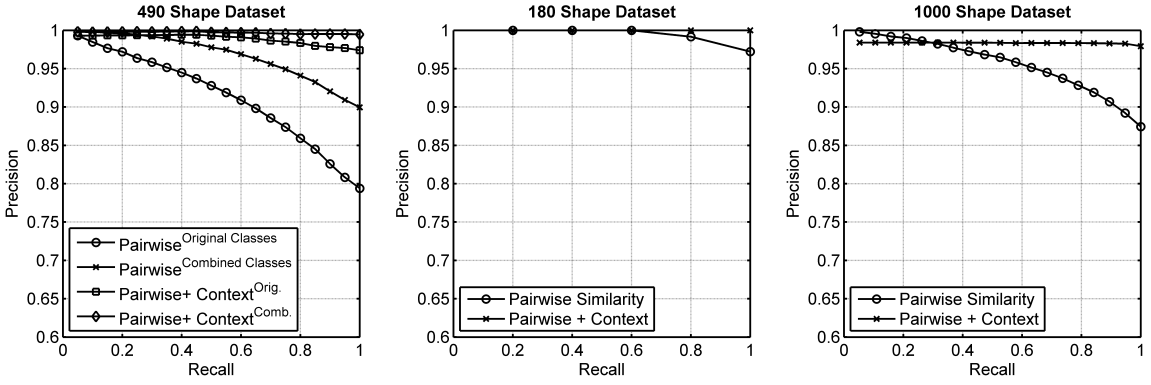


Figure 4.5: Precision Recall graphs for retrieval experiments on three datasets.

features method. Similarly, for the 180- and 1000-shapes datasets, both with and without using context information, favorable precision-recall results are obtained.

On the other hand, recently, there is an increased number of shape description methods based on eigenfunctions and/or eigenvalues of the Laplace-Beltrami operator computed over the shape silhouette [137, 59, 135, 140, 141]. However, it is known that Laplacian eigen-bases computed independently on shapes going through general visual changes are often incompatible with each other [144]. Therefore, approaches such as heat kernel signature [135] will not be able to perform well in a shape retrieval scenario with shapes of non-isometric changes. To overcome this shortcoming of inconsistency across shape variations, Laplacian eigenfunctions were computed jointly for multiple shapes in [144]. However, this approach requires re-computation of a joint Laplacian eigen-bases set between each pair of shapes along with pre-computed shape correspondences, which would constitute a limitation in a generic and practical shape retrieval application.

4.3 Conclusions

We presented a consistent and expressive shape representation, which is based on a 1-parameter family of shape field that is constructed by adding the ρ-dimension to the solution field of screened Poisson equation. The introduced increase in dimensions is followed by a reduction through projection onto principal modes of variation along the ρ-dimension. Contrasting to boundary-based approaches, this region-based representation allows the landmarks to be obtained from the whole shape domain, which leads to robustness to artifacts that can occur in shape boundaries. In addition, correspondences are obtained by combination of *almost local* (shape field measurements) and *global* (eigenvectors) cues due to the characteristic of the projections. The projections show consistency across various scenarios: different shape poses, deformations, occlusions and clutter.

Thanks to both the holistic and regional nature of the provided shape representation, a SIFT-based image matching framework, could be used in shape matching, for the first time to our knowledge. Thus, the SIFT descriptors constructed from the local feature frames detected over a given set of shape projections are utilized in shape matching. Moreover,

in addition to matching projection characteristics, in order to enforce geometrically likely configurations for 2D shapes, outliers are rejected through a planar homography assumption, which acts as a regularizer for the matching procedure. Even the surprisingly simple idea employed as the shape similarity measure, which is the total number of retained correspondences across corresponding shape projections of the two shapes being compared, achieves very good performance in matching as demonstrated by the performance over three common shape datasets. The presented shape matching scheme performs favorably among some popular shape retrieval methods.

5

Local Convexity Encoding Fields

In this chapter, a shape field that encodes convexity and concavity inside the shape domain is introduced. The proposed field is smooth, fluctuating(except for circle shape) and scale invariant. The distance properties of external parts are encoded explicitly, while isocontours in the central regions lay emphasis on concave regions near the boundary.

5.1 A new gaze into the hyper-field

When the whole variance in the field $v(\vec{x}, \rho)$ is considered, it should be noted that most of the variance can be accounted for the distances of nodes to the boundary of the shape, the source in the Poisson PDE. The appropriate example would be the case for a circle shape in $\mathbb{R}^2$, where $v(\vec{x}, \rho)\ R^3 \to \mathbb{R}$ can be expressed as $v_{Circle}(\mathcal{D}(\vec{x}), \rho)\ \mathbb{R}^2 \to \mathbb{R}$, since the field is constant on points equidistant to the boundary Γ:

$$v(\boldsymbol{x}\,, \rho) = v(\boldsymbol{x}\,, \rho) \qquad \{\forall(\boldsymbol{x}\,), (\boldsymbol{x}\,) \in \Omega \mid \mathcal{D}(\boldsymbol{x}\,) = \mathcal{D}(\boldsymbol{x}\,)\}\,, \tag{5.1}$$

where $\mathcal{D}$ is the Distance Transform for the shape. So, the field $v_{Circle}(\mathcal{D}(\vec{x}), \rho)$ does not contain any shape information other than a monotone relationship between values of the distance transform and v field. Of course this is not the case for the v field calculated on any shape domain. Motivated from this, it would be natural define the *shape information* preserved in the field as *deviations from the relationship between* $v(\boldsymbol{x}\,, \rho)$ *and* $\mathcal{D}(\vec{x})$.

Keeping in mind the definition above, consider the first iterations of diffusion procedure

($\rho = \infty$), the values of $v(\vec{x})$ will mimic an exponential function of distance to boundary of each node, similar to very low values of ρ as indicated in §3.3. This analogy holds until the propagating fronts of diffusion emanating from different sources meets at the local symmetry points. After this instance of the diffusion procedure for each symmetry point, the deviation described starts and encryption of valuable shape information begins.

Remarkably, the separation of the sources of variance of $log(v(\vec{x}, \rho))$ using PCA results in projections that are greatly in accordance with this description. A projection using the first eigenvector always leads to a field that encrypts distance relations w.r.t the boundary up to a scale in a smooth fashion. Whereas the second eigenvector projects the data such that the nodes near a local symmetry locus are distinguished from the rest of the nodes. The rest of the eigenvectors, which explain very little variance in comparison expose secondary details that rise from symmetry axis. Projections obtained using first two principal components of $log(v(\vec{x}, \rho))$ are given in Fig. 5.1.

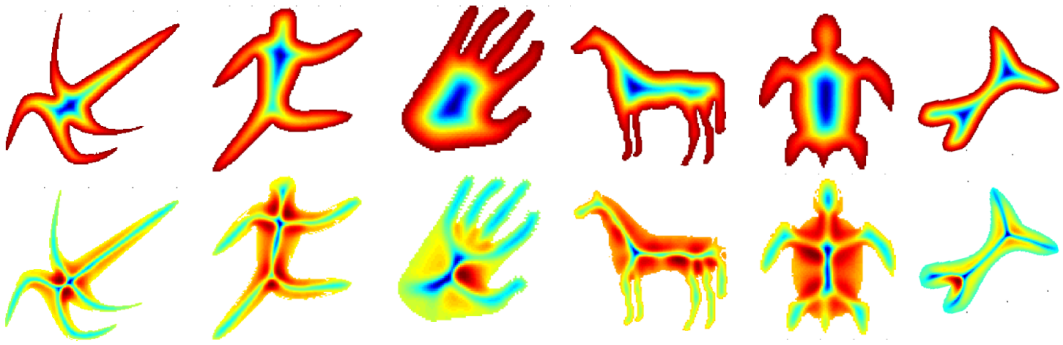


Figure 5.1: Projections using first two principal components of $log(v(x_2, y_2, \rho))$ on various shapes of closed contours.

As expected, distance to boundary constitutes the largest portion of the variance in the data. The fact that the shape information is separated by the second basis in this manner in the log domain confirms the consistency observed in the principal components of the original field. Additionaly, this point of view reinforces the relationship[85][126] established between screened Poisson PDE and Medial Axis Transform(MAT)[64].

5.2 Local convexity encoding fields

The numerical estimation of the deviations from the relationship between $v(\vec{x}, \rho)$ and $\mathcal{D}(\vec{x})$, is modeled as an answer to the following question for a node at point $(\vec{x})$:

> *Considering the distance of the node to boundary ($\mathcal{D}(\vec{x})$), how different would the intensity of node ($v(\vec{x}, \rho)$) be at time t if the source(shape boundary) was a perfect circle.*

The answer to this question for each node separately, leads to the difference of the field $v(\vec{x}, \rho, t)$ from a reference shape field($v_{Ref}(\mathcal{D}(\vec{x}), \rho, t)$), which is described on the shape domain and is a non-linear function of the distance transform for all ρ. For computational convenience we will do it for one rho (i.e. $\rho = \infty$). We calculate $v_{Ref} = v_{Circle}$ by solving the screened Poisson PDE for a circle given radius r and use the values of $v_{Circle}(\mathcal{D}(\vec{x}), \rho)$ as a look-up table. The ρ parameter and time step used should be the same as the parameters used while solving for $v(\vec{x}, \rho)$. The values of $v_{Ref}(\mathcal{D}(\vec{x}), t)$ are of course highly dependent on the radius of the circle. How the values of v_{Ref} changes in time for several radii are given in Fig. 5.2. Notice that the values for $\mathcal{D}(\vec{x}) = 15$ becomes much lower at any t, as radius of the reference circle increases. This is due to decreasing $\frac{perimeter}{area}$ ratio.

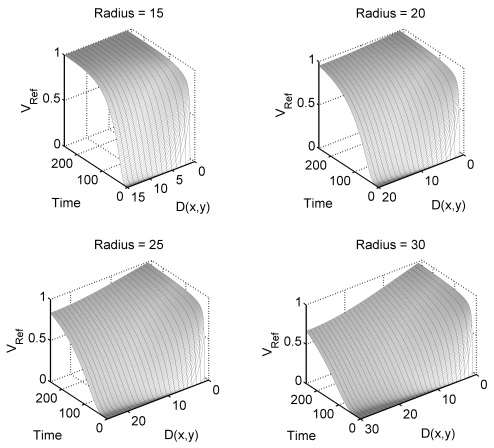


Figure 5.2: $V_{Ref}(\mathcal{D}(\vec{x}))$ vs time as reference radius increases.

We choose the maximum value of $\mathcal{D}(\vec{x})$ as the radius of the reference circle. With this choice, the difference of the Poisson PDE field and reference fields separates the external

parts and the core of the shape with its zero crossing. The resulting field at time t is formulated as:

$$\Psi(\vec{x}, t) = v_\rho(\vec{x}, t) - v_{Ref}(\mathcal{D}(\vec{x}), t) \tag{5.2}$$

The separation of parts using the field $\Psi(\vec{x}, t)$ is scale invariant. This is due to the selection of $max(\mathcal{D}(\vec{x}))$ as a radius to calculate V_{Circle}. Yet, the separation by the zero-crossing of the field alters in time, the invariance can be analysed as instances of time selected identical to the minimum of the v field. In Fig. 5.3 $\Psi(\vec{x}, t)$ for two cat shapes are presented along with separated parts.

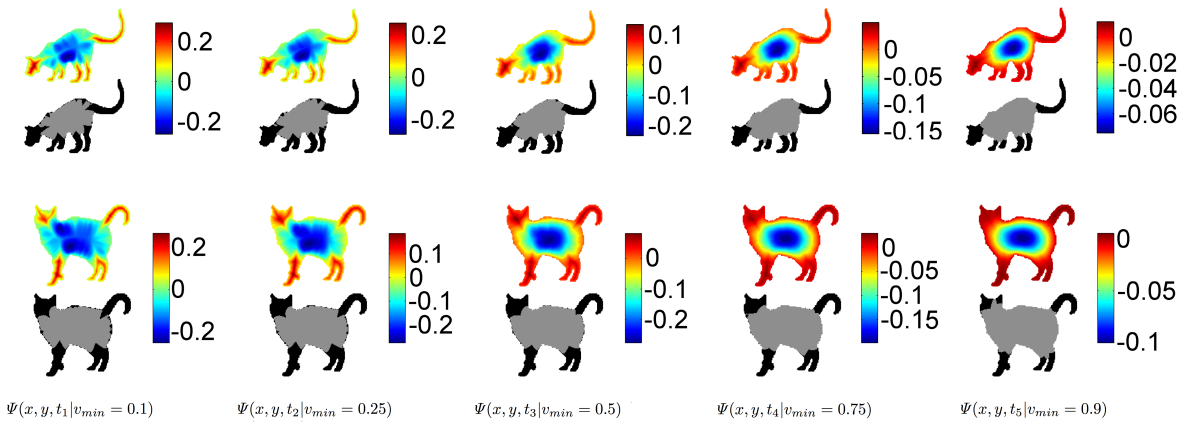


Figure 5.3: $\Psi(\vec{x}, t)$ and partitioning by the zero level curve are presented for two cat shapes in five time instances. Time instances are matched by the minimum vale of $v_\rho(\vec{x}, t)$. The values used are 0.1 for t_1, 0.25 for t_2, 0.5 for t_3, 0.75 for t_4 and 0.9 for t_5.

At low values of t, the field is not yet fully influenced by the boundary condition, and for very large values of t, the precision becomes problematic for both terms used for the calculation of $\Psi(\vec{x}, t)$ are very close to 1. In order to extract external parts in a robust manner, we use $\Psi(\vec{x}, t)$ as a source term for an inhomogeneous heat equation. The field for robust part extraction is formulated as:

$$\frac{\partial A(\vec{x}, t)}{\partial t} = \alpha \left(\sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} A(\vec{x}, t) \right) + f(\Psi(\vec{x}, t)) \tag{5.3}$$

where α is the factor to control the influence of the diffusion term. We use the function f to regularize $\Psi(\vec{x},t)$, such that the influence of $\Psi(\vec{x},t)$ as a source will be lowered if the minimum value of $v(\vec{x},t)$(can be considered as approximation to the shape center) is very low or very high. This is accomplished by multiplication with a factor that is normally distributed with a mean of 0.5. f is formulated below, where σ should be specified a priori.

$$f(\Psi(\vec{x},t)) = \Psi(\vec{x},t)e^{\frac{-(min(v(\vec{x},t))-0.5)^2}{2\sigma^2}} \tag{5.4}$$

The field is solved by a finite difference method using Neumann boundary condition. The solution, $A_{final}(\vec{x})$ is $A(\vec{x},t_{final})$, where t_{final} corresponds to the time of convergence for $v_\rho(\vec{x},t)$. A_{final} is the best harmonic approximation of the hyper-field $f(\Psi(\vec{x},t))$. The steady state of the inhomogeneous heat equation can be expressed as the sum of the source term, $\Psi(\vec{x},t)$'s on considered t instances, with more influence on the instances where $min(v_\rho(\vec{x},t))$ is close to 0.5. This is the region where $\Psi(\vec{x},t)$ is more informative (Fig. 5.2) and less affected by precision related problems. $A_{final}(\vec{x})$ fields and extracted parts for various shapes are presented in Fig. 5.4.

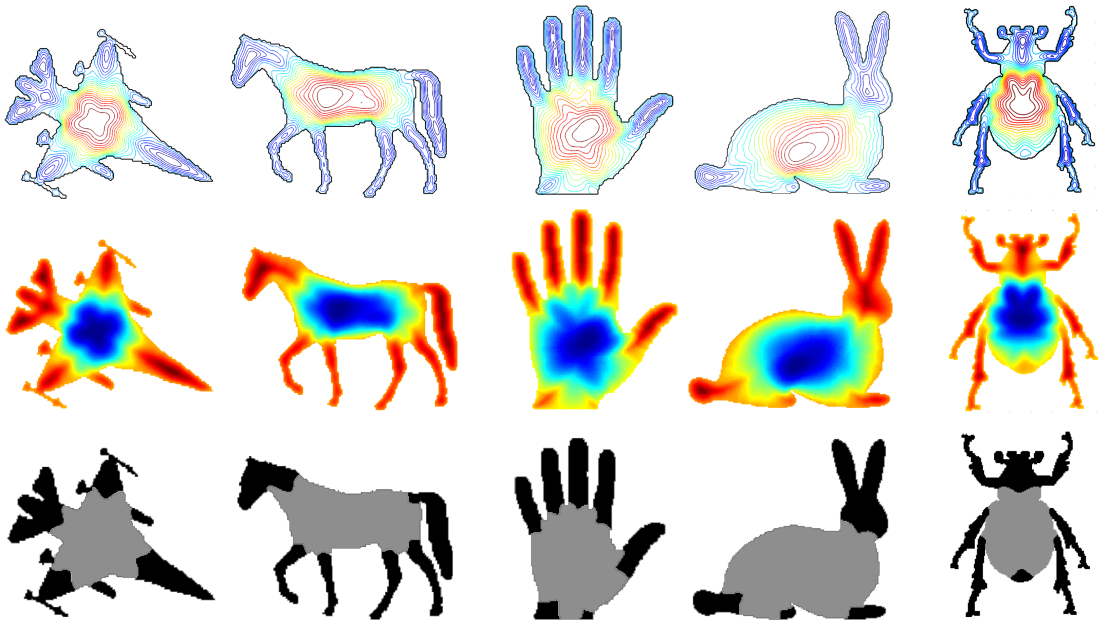


Figure 5.4: $A_{final}(\vec{x})$ fields and extracted parts for various shapes.

The parameters that should be considered for part extraction are the standard deviation σ in the robustness function f and α. The effect of the σ parameter is not affected by scale also, since it takes values with respect to v_{min}. The fields for two shapes will be consistent as long as the same value σ and α is used for calculating both fields. For obtaining the results in Fig.5.4, we used $\sigma = 0.5$ and $\alpha = 0.005$.

Gorelick et.al. [52] also proposed a method for part extraction, which requires the thresholding of an all positive field. This method would give very unstable results for shapes with protrusions of different kinds. Although our method results into a field whose level curves are slightly dependent on α and σ, the separation of the variables do not get affected from these parameters. In addition while Gorelick's approach would require different thresholds for different scales even for the same shape, the separation in our field is robust to scale changes.

The distinction has a hierarchical aspect also. Considering the intensities of the resulting fields, the parts that have higher maximum intensities can be considered relatively significant compared to other parts. For example, in Fig.9, the front and the rear parts of the F-16 are more significant compared to wings, head and tail of the horse is more significant compared to legs, fingers are more significant compared to parts of the wrist that is cut, the head and ears of the bunny can be separated and are more significant compared to legs, head of the bug is more significant compared to four legs. In this description, significance implies a larger region that is affected by symmetry points and not affected by concavities(will be clearer in the following), which is obviously not the only way to form a hierarchy.

The factors that cause the distinction of parts in the field are the (i)local symmetry points, which tend to have larger values locally, and (ii)concavities, which tend to have lower values locally (compared to a circular boundary at same distance). This fact can be observed at Fig.5.5, where $A(\vec{x}, t)$ is projected via the process in §3.4.2, replacing the hyper-field $v(\vec{x}, \rho)$ with $A(\vec{x}, t)$. The hyper-field $A(\vec{x}, t)$ is obtained by measurements sampled from the time domain as field A is evolving. The first principal direction leads to a projection that is very similar to A field. The second principal component demonstrates the race between symmetry points and concavities.

The detection of the concavities is also proposed by Gorelick et.al. [52] but the results also include the concavities in the external parts, additionally the detection depends on a soft threshold parameter. In our method, the concavities detected are regularized by the effect of the local symmetry points. The detection is also affected by the global distribution of concave points (not spatially but as the number of concave points and intensities because of the way we use PCA). This framework would allow the use of the intensities of the concave points to include a hierarchical discrimination. For example the concavities separating the head of the bug in Fig.5.5 are much significant in comparison to concave points that separate legs. Also the leaf shape (from MPEG-7 database) can be analyzed hierarchically. Two leaves stem from the same base which stems from a circular base. The concave points in the connection of two smaller leaves are less significant compared to the concavities that connect the larger segments near the circular base.

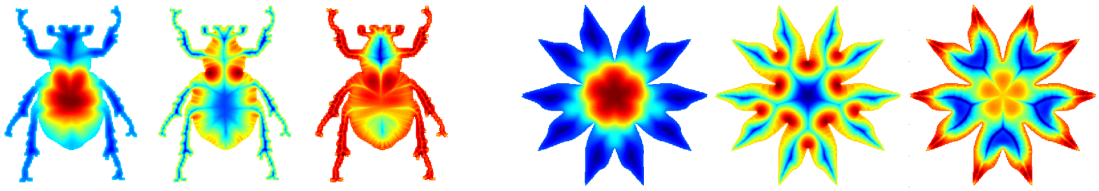


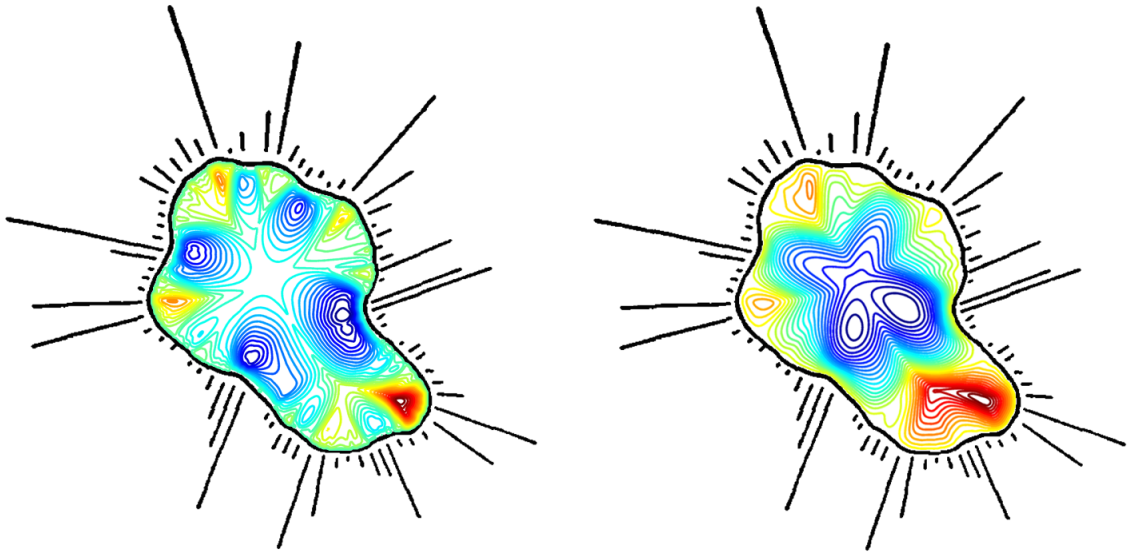
Figure 5.5: Projections using the principal components of the hyper-field $A(\vec{x}, t)$ in the time domain for two shapes.

5.3 Discussion and Conclusion

The proposed shape field, A, is mainly motivated from the continuum of fields described in Chapter 3, specifically it is modeled as the deviation from the circle (or sphere for 3d) case, where the potential in any node inside the shape domain can be parameterized by a single variable: radius. When the radius is selected as the maximum distance to boundary, the deviations encode external parts explicitly.

In his seminal work, *Some Informational Aspects of Visual Perception,1954* [2], Fred Attneave performed psychological experiments to demonstrate that visual data is highly redundant and portrays that points of high curvature on shape boundaries are informative

and perceptually relevant. The famous cat experiment was created by abstracting 38 points of maximum curvature from the contours of a sleeping cat, and connecting these points appropriately with a straightedge. Even the very simple representation of the cat could be reliably recognized in the drawing. It has been suggested that such points have high information content, which was a motivation for earlier methods in partitioning shape boundaries based on high curvature [212, 213, 214, 215, 216]. Another experiment, perhaps one that did not recieve as much attention as the cat experiment, involves eighty subjects that were instructed to draw, a pattern of 10 dots which would resemble the shape of closed contours as closely as possible. The local convexity encoding field is calculated for the shape used in this experiment to demonstrate how the proposed regional representation respects the perceptually important cues on the shape boundary in Fig.5.6.



(a) $A(\boldsymbol{x}, t|v_{min} = 0.1)$ (b) $A_{final}(\boldsymbol{x})$

Figure 5.6: $A(\vec{x}, t)$ at two different t values, for the shape presented in Attneave's work[2]. Subjects attempted to approximate the closed figure shown above with a pattern of 10 dots. Radiating bars indicate the relative frequency with which various portions of the outline were represented by dots chosen. Note how the A hyperfield simulates the human subjects in marking convex and concave points on the boundary of the shape.

Attneave's findings:

It is clear that subjects show a great deal of agreement in their abstractions

> of points best representing the shape, and most of these points are taken from regions where the contour is most different from a straight line. This conclusion is verified by detailed comparisons of dot frequencies with measured curvatures on both the figure shown and others.

were greatly in accordance with Hebb's observations[31], that the regions on the boundary that have high curvature are informative: *information is further concentrated at points where a contour changes direction most rapidly.* The locally convexity encoding field, as shown in Fig.5.6(a), is highly related to those points. In fact, the negative and positive intensities are almost *leaking* inside the shape from those significant points and as time goes by, an *agreement* of those are declared as the smooth shape representation: $A_{final}(\boldsymbol{x})$. The field provides very intuitive results as a non-parametric external part extraction, which is consistent across pose changes. Note that consistency claim is currently solely observational. Moreover, the rich local information encoded is completely novel and can be considered promising to be utilized in a variety shape analysis applications.

6

Conclusions and Future Directions

To sum up, the foundations of the new scalable fluctuating distance fields, shape hyper-fields and the local convexity encoding fields are introduced. The strength of the methods is demonstrated in numerous applications:

- Interactive tumor protrusion segmentation via fluctuating distance fields
- Tumor follow-up registration based on constancy assumption of fluctuating distance fields as underlying shape representation.
- Regional shape partitioning and boundary decomposition using non-negative sparse coding to screened Poisson hyper-fields.
- Planar shape classification using primitive moments of nodal structures of SPEM's.
- 3D Shape retrieval using VLAD descriptors applied to SPEM's.
- Planar shape matching and retrieval using SIFT descriptors.
- Shape partitioning and representation using local convexity encoding fields.

Representations presented in this thesis are constructed over the whole shape domain, therefore, it is certainly possible to either define new shape measures or adopt existing popular descriptors. We believe that they will find ubiquitous use in various shape analysis applications.

For future directions of research, invariance properties of the representations will be investigated in more detail. The non-rigid retrieval experiments reveal that the SPEM representation for instance is highly robust for encoding articulated motion. Such a property could very well be undesired in some applications, for instance modelling of anatomical structures. We will investigate spatially varying conditioning methods for random walkers in order to obtain an isometry-varying representation. The spatial variation of conditioning could be obtained by curvature or surface properties. An alternative direction of research is compositing the shape fields with other available measurements from inside the shape domain to address problems of object characterization.

Bibliography

[1] S. Tari and M. Genctav, "From a non-local ambrosio-tortorelli phase field to a randomized part hierarchy tree," *Journal of Mathematical Imaging and Vision*, pp. 1–18, 2013.

[2] F. Attneave, "Some informational aspects of visual perception." *Psychological review*, vol. 61, no. 3, p. 183, 1954.

[3] R. Kimmel, *Numerical geometry of images: Theory, algorithms, and applications*. Springer, 2004.

[4] A. Srivastava, S. H. Joshi, W. Mio, and X. Liu, "Statistical shape analysis: Clustering, learning, and testing," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 4, pp. 590–602, 2005.

[5] Y. Gao, T. Riklin-Raviv, and S. Bouix, "Shape analysis, a field in need of careful validation," *Human brain mapping*, 2014.

[6] G. Gerig, M. Styner, D. Jones, D. Weinberger, and J. Lieberman, "Shape analysis of brain ventricles using spharm," in *Mathematical Methods in Biomedical Image Analysis, 2001. MMBIA 2001. IEEE Workshop on*. IEEE, 2001, pp. 171–178.

[7] M. Styner, G. Gerig, J. Lieberman, D. Jones, and D. Weinberger, "Statistical shape analysis of neuroanatomical structures based on medial models," *Medical Image Analysis*, vol. 7, no. 3, pp. 207–220, 2003.

[8] M. Styner, J. A. Lieberman, D. Pantazis, and G. Gerig, "Boundary and medial shape analysis of the hippocampus in schizophrenia," *Medical image analysis*, vol. 8, no. 3, pp. 197–203, 2004.

[9] M. K. Chung, K. M. Dalton, and R. J. Davidson, "Tensor-based cortical surface morphometry via weighted spherical harmonic representation," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 8, pp. 1143–1151, 2008.

[10] A. Qiu, D. Crocetti, M. Adler, E. Mahone, M. Denckla, M. Miller, and S. Mostofsky, "Basal ganglia volume and shape in children with attention deficit hyperactivity disorder," *American Journal of Psychiatry*, vol. 166, no. 1, pp. 74–82, 2009.

[11] L. Younes, J. T. Ratnanather, T. Brown, E. Aylward, P. Nopoulos, H. Johnson, V. A. Magnotta, J. S. Paulsen, R. L. Margolis, R. L. Albin *et al.*, "Regionally selective atrophy of subcortical structures in prodromal hd as revealed by statistical shape analysis," *Human brain mapping*, 2012.

[12] S.-G. Kim, M. K. Chung, S. M. Schaefer, C. Van Reekum, and R. J. Davidson, "Sparse shape representation using the laplace-beltrami eigenfunctions and its application to modeling subcortical structures," in *Mathematical Methods in Biomedical Image Analysis (MMBIA), 2012 IEEE Workshop on*. IEEE, 2012, pp. 25–32.

[13] X. Yang, A. Goh, S.-H. A. Chen, and A. Qiu, "Evolution of hippocampal shapes across the human lifespan," *Human brain mapping*, vol. 34, no. 11, pp. 3075–3085, 2013.

[14] J. Joseph, C. Warton, S. W. Jacobson, J. L. Jacobson, C. D. Molteno, A. Eicher, P. Marais, O. R. Phillips, K. L. Narr, and E. M. Meintjes, "Three-dimensional surface deformation-based shape analysis of hippocampus and caudate nucleus in children with fetal alcohol spectrum disorders," *Human brain mapping*, vol. 35, no. 2, pp. 659–672, 2014.

[15] A. Shamir, "A survey on mesh segmentation techniques," in *Computer graphics forum*, vol. 27. Wiley Online Library, 2008, pp. 1539–1556.

[16] A. M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov, "Shape google: Geometric words and expressions for invariant shape retrieval," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 1, p. 1, 2011.

[17] O. Van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or, "A survey on shape correspondence," in *Computer Graphics Forum*, vol. 30, no. 6. Wiley Online Library, 2011, pp. 1681–1707.

[18] L. Shen, H. Farid, and M. A. McPeek, "Modeling three-dimensional morphological structures using spherical harmonics," *Evolution*, vol. 63, no. 4, pp. 1003–1016, 2009.

[19] C. Costa, F. Antonucci, F. Pallottino, J. Aguzzi, D.-W. Sun, and P. Menesatti, "Shape analysis of agricultural products: a review of recent research advances and potential application to computer vision," *Food and Bioprocess Technology*, vol. 4, no. 5, pp. 673–692, 2011.

[20] H. Ziezold, "Mean figures and mean shapes applied to biological figure and shape distributions in the plane," *Biometrical journal*, vol. 36, no. 4, pp. 491–510, 1994.

[21] C. G. Small, *The Statistical Theory of Shape, Springer Series in Statistics*. New York: Springer-Verlag, 1996.

[22] I. L. Dryden and K. V. Mardia, *Statistical shape analysis*. John Wiley & Sons New York, 1998, vol. 4.

[23] H. Krim and A. J. Yezzi, *Statistics and analysis of shapes*. Springer, 2006.

[24] D. G. Kendall, D. Barden, T. Carne, and H. Le, *Shape and shape theory*. John Wiley & Sons, 2009, vol. 500.

[25] A. Pasupathy and C. E. Connor, "Population coding of shape in area v4," *Nature neuroscience*, vol. 5, no. 12, pp. 1332–1338, 2002.

[26] Z. Pylyshyn, "Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception," *Behavioral and brain sciences*, vol. 22, no. 03, pp. 341–365, 1999.

[27] S. Loncaric, "A survey of shape analysis techniques," *Pattern recognition*, vol. 31, no. 8, pp. 983–1001, 1998.

[28] M. Wertheimer, "Laws of organization in perceptual forms," *A source book of Gestalt psychology*, pp. 71–88, 1938.

[29] W. D. Ellis, *A source book of Gestalt psychology*. Psychology Press, 1999, vol. 2.

[30] K. Koffka, *Principles of Gestalt psychology*. Routledge, 2013.

[31] D. Hebb, *0.(1949) The organization of behavior*. Wiley, New York, 1968.

[32] J. J. Gibson, "The perception of the visual world." 1950.

[33] D. Marr, "Early processing of visual information," *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, vol. 275, no. 942, pp. 483–519, 1976.

[34] ——, "A theory for cerebral neocortex," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, pp. 161–234, 1970.

[35] B. K. Horn, "Height and gradient from shading," *International journal of computer vision*, vol. 5, no. 1, pp. 37–75, 1990.

[36] Q. Zheng and R. Chellappa, "Estimation of illuminant direction, albedo, and shape from shading," in *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*. IEEE, 1991, pp. 540–545.

[37] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah, "Shape-from-shading: a survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, no. 8, pp. 690–706, 1999.

[38] W. Hoff and N. Ahuja, "Surfaces from stereo: Integrating feature matching, disparity estimation, and contour detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 2, pp. 121–136, 1989.

[39] J. Aloimonos, "Shape from texture," *Biological cybernetics*, vol. 58, no. 5, pp. 345–360, 1988.

[40] R. Horaud and M. Brady, "On the geometric interpretation of image contours," *Artificial intelligence*, vol. 37, no. 1, pp. 333–353, 1988.

[41] S. K. Nayar and Y. Nakagawa, "Shape from focus," *Pattern analysis and machine intelligence, IEEE Transactions on*, vol. 16, no. 8, pp. 824–831, 1994.

[42] D. G. Lowe, "Three-dimensional object recognition from single two-dimensional images," *Artificial intelligence*, vol. 31, no. 3, pp. 355–395, 1987.

[43] J. J. Koenderink, "The structure of images," *Biological cybernetics*, vol. 50, no. 5, pp. 363–370, 1984.

[44] J. J. Koenderink and A. J. van Doorn, "Dynamic shape," *Biological cybernetics*, vol. 53, no. 6, pp. 383–396, 1986.

[45] D. G. Kendall, "The diffusion of shape," *Advances in applied probability*, pp. 428–430, 1977.

[46] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker, "Shock graphs and shape matching," *International Journal of Computer Vision*, vol. 35, no. 1, pp. 13–32, 1999.

[47] L. J. Latecki and R. Lakamper, "Shape similarity measure based on correspondence of visual parts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 10, pp. 1185–1190, 2000.

[48] B. Leibe and B. Schiele, "Analyzing appearance and contour based methods for object categorization," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 2. IEEE, 2003, pp. II–409.

[49] Z. Tu and A. L. Yuille, "Shape matching and recognition–using generative models and informative features," in *Computer Vision-ECCV 2004*. Springer, 2004, pp. 195–209.

[50] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing their shock graphs," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 5, pp. 550–571, 2004.

[51] A. Shokoufandeh, D. Macrini, S. Dickinson, K. Siddiqi, and S. W. Zucker, "Indexing hierarchical structures using graph spectra," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 7, pp. 1125–1140, 2005.

[52] L. Gorelick, M. Galun, E. Sharon, R. Basri, and A. Brandt, "Shape representation and classification using the poisson equation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 12, pp. 1991–2005, 2006.

[53] G. McNeill and S. Vijayakumar, "Hierarchical procrustes matching for shape retrieval," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1. IEEE, 2006, pp. 885–894.

[54] P. F. Felzenszwalb and J. D. Schwartz, "Hierarchical matching of deformable shapes," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on.* IEEE, 2007, pp. 1–8.

[55] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 2, pp. 286–299, 2007.

[56] D. M. Gavrila, "A bayesian, exemplar-based approach to hierarchical shape matching," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 8, pp. 1408–1421, 2007.

[57] X. Bai and L. J. Latecki, "Path similarity skeleton graph matching," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 7, pp. 1282–1292, 2008.

[58] E. Konukoglu, B. Glocker, A. Criminisi, and K. Pohl, "Wesd-weighted spectral distance for measuring shape dissimilarity," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 9, pp. 2284–2297, 2013.

[59] M. Reuter, F.-E. Wolter, and N. Peinecke, "Laplace–beltrami spectra as shape-dna of surfaces and solids," *Computer-Aided Design*, vol. 38, no. 4, pp. 342–366, 2006.

[60] E. Konukoglu, B. Glocker, D. H. Ye, A. Criminisi, and K. M. Pohl, "Discriminative segmentation-based evaluation through shape dissimilarity," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 12, pp. 2278–2289, 2012.

[61] T. Pavlidis, "Algorithms for shape analysis of contours and waveforms." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, no. 4, pp. 301–312, 1980.

[62] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern recognition*, vol. 37, no. 1, pp. 1–19, 2004.

[63] M. Yang, K. Kpalma, J. Ronsin *et al.*, "A survey of shape feature extraction techniques," *Pattern recognition*, pp. 43–90, 2008.

[64] H. Blum *et al.*, "A transformation for extracting new descriptors of shape," *Models for the perception of speech and visual form*, vol. 19, no. 5, pp. 362–380, 1967.

[65] H. Blum, "Biological shape and visual science (part i)," *Journal of theoretical Biology*, vol. 38, no. 2, pp. 205–287, 1973.

[66] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker, "Shapes, shocks, and deformations i: the components of two-dimensional shape and the reaction-diffusion space," *International journal of computer vision*, vol. 15, no. 3, pp. 189–224, 1995.

[67] S. Sclaroff and A. Pentland, "Generalized implicit functions for computer graphics," *SIGGRAPH Comput. Graph.*, vol. 25, no. 4, pp. 247–250, Jul. 1991. [Online]. Available: http://doi.acm.org/10.1145/127719.122745

[68] V. V. Savchenko, A. A. Pasko, O. G. Okunev, and T. L. Kunii, "Function representation of solids reconstructed from scattered surface points and contours," in *Computer Graphics Forum*, vol. 14. Wiley Online Library, 1995, pp. 181–188.

[69] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations," *J. Comput. Phys.*, vol. 79, no. 1, pp. 12–49, 1988.

[70] J. A. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, 2nd ed., ser. Cambridge Monographs on Applied and Computational Mathematics. Cambridge: Cambridge University Press, 1999, vol. 3.

[71] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 1. IEEE, 2000, pp. 316–323.

[72] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 17, no. 2, pp. 158–175, 1995.

[73] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International journal of computer vision*, vol. 22, no. 1, pp. 61–79, 1997.

[74] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Gradient flows and geometric active contour models," in *Computer Vision, 1995. Proceedings., Fifth International Conference on.* IEEE, 1995, pp. 810–815.

[75] T. F. Chan and L. A. Vese, "Active contours without edges," *Image Processing, IEEE Transactions on*, vol. 10, no. 2, pp. 266–277, 2001.

[76] A. Tsai, A. Yezzi Jr, W. Wells III, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "Model-based curve evolution technique for image segmentation," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1. IEEE, 2001, pp. I–463.

[77] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *International Journal of Computer Vision*, vol. 46, no. 3, pp. 223–247, 2002.

[78] M. Rousson, T. Brox, and R. Deriche, "Active unsupervised texture segmentation on a diffusion based feature space," in *Computer vision and pattern recognition, 2003. Proceedings. 2003 IEEE computer society conference on*, vol. 2. IEEE, 2003, pp. II–699.

[79] M. Heiler and C. Schnörr, "Natural image statistics for natural image segmentation," *International Journal of Computer Vision*, vol. 63, no. 1, pp. 5–19, 2005.

[80] T. Brox and J. Weickert, "A tv flow based local scale measure for texture discrimination," in *Computer Vision-ECCV 2004*. Springer, 2004, pp. 578–590.

[81] G. Unal, A. Yezzi, and H. Krim, "Information-theoretic active polygons for unsupervised texture segmentation," *International Journal of Computer Vision*, vol. 62, no. 3, pp. 199–220, 2005.

[82] D. Cremers and S. Soatto, "Motion competition: A variational approach to piecewise parametric motion segmentation," *International Journal of Computer Vision*, vol. 62, no. 3, pp. 249–265, 2005.

[83] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for motion estimation and tracking," *Computer Vision and Image Understanding*, vol. 97, no. 3, pp. 259–282, 2005.

[84] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *International journal of computer vision*, vol. 72, no. 2, pp. 195–215, 2007.

[85] Z. Tari, J. Shah, and H. Pien, "Extraction of shape skeletons from grayscale images," *Computer Vision and Image Understanding*, vol. 66, no. 2, pp. 133–146, 1997.

[86] P. O. Hoyer, "Non-negative sparse coding," in *Neural Networks for Signal Processing, 2002. Proceedings of the 2002 12th IEEE Workshop on*. IEEE, 2002, pp. 557–565.

[87] Z. Lian, A. Godil, B. Bustos, M. Daoudi, J. Hermans, S. Kawamura, Y. Kurita, G. Lavoué, H. Van Nguyen, R. Ohbuchi *et al.*, "Shrec'11 track: Shape retrieval on non-rigid 3d watertight meshes." *3DOR*, vol. 11, pp. 79–88, 2011.

[88] H. Jégou, M. Douze, C. Schmid, and P. Pérez, "Aggregating local descriptors into a compact image representation," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 3304–3311.

[89] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2. Ieee, 1999, pp. 1150–1157.

[90] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[91] E. Konukoglu, X. Pennec, O. Clatz, and N. Ayache, "Tumor growth modeling in oncological image analysis," in *Handbook of Medical Image Processing and Analysis - New edition*, I. Bankman, Ed. Academic Press, December 2008, ch. 18, pp. 297–307.

[92] R. Tombropoulos, A. Schweikard, J.-C. Latombe, and J. Adler, "Treatment planning for image-guided robotic radiosurgery," in *Computer Vision, Virtual Reality and Robotics in Medicine*. Springer, 1995, pp. 131–137.

[93] F. M. Khan, *The physics of radiation therapy*. Lippincott Williams & Wilkins, 2009.

[94] I. Biederman, "Recognition-by-components: a theory of human image understanding." *Psychological review*, vol. 94, p. 115, 1987.

[95] D. D. Hoffman and W. A. Richards, "Parts of recognition," *Cognition*, vol. 18, pp. 65–96, 1984.

[96] X. Chen, A. Golovinskiy, and T. Funkhouser, "A benchmark for 3d mesh segmentation," in *ACM Transactions on Graphics (TOG)*, vol. 28, no. 3. ACM, 2009, p. 73.

[97] A. Golovinskiy and T. Funkhouser, "Consistent segmentation of 3d models," *Computers & Graphics*, vol. 33, no. 3, pp. 262–269, 2009.

[98] E. Kalogerakis, A. Hertzmann, and K. Singh, "Learning 3d mesh segmentation and labeling," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 4, p. 102, 2010.

[99] M. Attene, S. Katz, M. Mortara, G. Patané, M. Spagnuolo, and A. Tal, "Mesh segmentation-a comparative study," in *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on.* IEEE, 2006, pp. 7–7.

[100] M. P. Do Carmo and M. P. Do Carmo, *Differential geometry of curves and surfaces.* Prentice-Hall Englewood Cliffs, 1976, vol. 2.

[101] H. Pottmann, T. Steiner, M. Hofer, C. Haider, and A. Hanbury, "The isophotic metric and its application to feature sensitive morphology on surfaces," *Computer Vision-ECCV 2004*, pp. 18–23, 2004.

[102] Y.-K. Lai, Q.-Y. Zhou, S.-M. Hu, J. Wallner, D. Pottmann *et al.*, "Robust feature classification and editing," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 13, no. 1, pp. 34–45, 2007.

[103] K. Crane, C. Weischedel, and M. Wardetzky, "Geodesics in heat: A new approach to computing distance based on heat flow," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 5, p. 152, 2013.

[104] F. De Goes, S. Goldenstein, and L. Velho, "A hierarchical segmentation of articulated bodies," in *Computer graphics forum*, vol. 27, no. 5. Wiley Online Library, 2008, pp. 1349–1356.

[105] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and computational harmonic analysis*, vol. 21, no. 1, pp. 5–30, 2006.

[106] R. Liu, H. Zhang, A. Shamir, and D. Cohen-Or, "A part-aware surface metric for shape analysis," in *Computer Graphics Forum*, vol. 28, no. 2. Wiley Online Library, 2009, pp. 397–406.

[107] S. M. Pizer, D. S. Fritsch, P. A. Yushkevich, V. E. Johnson, and E. L. Chaney, "Segmentation, registration, and measurement of shape variation via image object shape," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 10, pp. 851–865, 1999.

[108] S. Joshi, S. Pizer, P. T. Fletcher, P. Yushkevich, A. Thall, and J. Marron, "Multiscale deformable model segmentation and statistical shape analysis using medial descriptions," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 5, pp. 538–550, 2002.

[109] M. Brady and H. Asada, "Smoothed local symmetries and their implementation," *The International Journal of Robotics Research*, vol. 3, no. 3, pp. 36–61, 1984.

[110] D. Shaked and A. M. Bruckstein, "Pruning medial axes," *Computer vision and image understanding*, vol. 69, no. 2, pp. 156–169, 1998.

[111] H. Blum and R. N. Nagel, "Shape description using weighted symmetric axis features," *Pattern recognition*, vol. 10, no. 3, pp. 167–180, 1978.

[112] J. August, K. Siddiqi, and S. W. Zucker, "Ligature instabilities in the perceptual organization of shape," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 2. IEEE, 1999.

[113] X. Mi and D. DeCarlo, "Separating parts from 2d shapes using relatability," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on.* IEEE, 2007, pp. 1–8.

[114] J. Feldman and M. Singh, "Bayesian estimation of the shape skeleton," *Proceedings of the National Academy of Sciences*, vol. 103, no. 47, pp. 18 014–18 019, 2006.

[115] C. A. Burbeck and S. M. Pizer, "Object representation by cores: Identifying and representing primitive spatial regions," *Vision research*, vol. 35, no. 13, pp. 1917–1930, 1995.

[116] D. Macrini, S. Dickinson, D. Fleet, and K. Siddiqi, "Bone graphs: Medial shape parsing and abstraction," *Computer Vision and Image Understanding*, vol. 115, no. 7, pp. 1044–1061, 2011.

[117] S. Tari, "Fluctuating distance fields, parts, three-partite skeletons," in *Innovations for Shape Analysis.* Springer, 2013, pp. 439–466.

[118] N. Paragios, M. Rousson, and V. Ramesh, "Non-rigid registration using distance functions," *Computer Vision and Image Understanding*, vol. 89, no. 2, pp. 142–165, 2003.

[119] F. Meyer, "Topographic distance and watershed lines," *Signal processing*, vol. 38, no. 1, 1994.

[120] A. Hamamci, N. Kucuk, K. Karaman, K. Engin, and G. Unal, "Tumor-cut: Segmentation of brain tumors on contrast enhanced mr images for radiosurgery applications," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 3, pp. 790–804, 2012.

[121] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," *Computer Vision-ECCV 2004*, pp. 25–36, 2004.

[122] B. Horn and B. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.

[123] L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 12, pp. 2247–2253, 2007.

[124] H. Haidar, S. Bouix, J. J. Levitt, R. W. McCarley, M. E. Shenton, and J. S. Soul, "Characterizing the shape of anatomical structures with poisson's equation," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 10, pp. 1249–1257, 2006.

[125] G. Aubert and J.-F. Aujol, "Poisson skeleton revisited: A new mathematical perspective," *Journal of Mathematical Imaging and Vision*, vol. 48, no. 1, pp. 149–159, 2014.

[126] C. Aslan, A. Erdem, E. Erdem, and S. Tari, "Disconnected skeleton: shape at its absolute scale," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 12, pp. 2188–2203, 2008.

[127] K. S. Gurumoorthy and A. Rangarajan, "Distance transform gradient density estimation using the stationary phase approximation," *SIAM J. Math. Analysis*, vol. 44, no. 6, pp. 4250–4273, 2012.

[128] M. Sethi, A. Rangarajan, and K. Gurumoorthy, "The schrödinger distance transform (sdt) for point-sets and curves," in *Computer Vision and Pattern Recognition, 2012 IEEE Conference on*. IEEE, 2012, pp. 198–205.

[129] A. Duci, A. J. Yezzi, S. K. Mitter, and S. Soatto, "Shape representation via harmonic embedding," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 656–662.

[130] A. Duci, A. Yezzi, S. Soatto, and K. Rocha, "Harmonic embeddings for linear shape analysis," *Journal of Mathematical Imaging and Vision*, vol. 25, no. 3, pp. 341–352, 2006.

[131] A. G. Belyaev, P.-A. Fayolle, and A. A. Pasko, "Signed Lp-distance fields." *Computer-Aided Design*, vol. 45, no. 2, pp. 523–528, 2013. [Online]. Available: http://dblp.uni-trier.de/db/journals/cad/cad45.html#BelyaevFP13

[132] P. Bhat, B. Curless, M. Cohen, and C. Zitnick, "Fourier analysis of the 2d screened poisson equation for gradient domain problems," in *Proc. 10th European Conf Computer Vision: ECCV*, 2008, pp. 114–128.

[133] M. Chuang and M. Kazhdan, "Interactive and anisotropic geometry processing using the screened poisson equation," *ACM Trans Graphics (TOG)*, vol. 30, no. 4, pp. 57–57, 2011.

[134] M. Kazhdan and H. Hoppe, "Screened poisson surface reconstruction," *ACM Trans Graphics (TOG)*, vol. 32, no. 3, pp. 29–29, 2013.

[135] J. Sun, M. Ovsjanikov, and L. Guibas, "A concise and provably informative multi-scale signature based on heat diffusion," in *Computer Graphics Forum*, vol. 28. Wiley Online Library, 2009, pp. 1383–1392.

[136] R. M. Rustamov, "Laplace-beltrami eigenfunctions for deformation invariant shape representation," in *Proceedings of the fifth Eurographics symposium on Geometry processing*. Eurographics Association, 2007, pp. 225–233.

[137] S. S. Lafon, "Diffusion maps and geometric harmonics," Ph.D. dissertation, Yale University, 2004.

[138] Y.-S. Liu, Q. Li, G.-Q. Zheng, K. Ramani, and W. Benjamin, "Using diffusion distances for flexible molecular shape comparison," *BMC Bioinformatics*, vol. 11, p. 480, 2010.

[139] M. M. Bronstein and I. Kokkinos, "Scale-invariant heat kernel signatures for non-rigid shape recognition," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 1704–1711.

[140] D. Raviv, M. M. Bronstein, A. M. Bronstein, and R. Kimmel, "Volumetric heat kernel signatures," in *Proceedings of the ACM workshop on 3D object retrieval*. ACM, 2010, pp. 39–44.

[141] M. Aubry, U. Schlickewei, and D. Cremers, "The wave kernel signature: A quantum mechanical approach to shape analysis," in *Computer Vision Workshops, 2011 IEEE International Conference on*. IEEE, 2011, pp. 1626–1633.

[142] M. Ovsjanikov, M.B-Chen, J. Solomon, A. Butscher, and L. Guibas, "Functional maps: A flexible representation of maps between shapes," *ACM Trans on Graphics (TOG)*, vol. 31, no. 4, pp. 30–30, 2013.

[143] M. Ovsjanikov, Q. Merigot, V. Patraucean, and L. Guibas, "Shape matching via quotient spaces," *Comp Graphics Forum*, vol. 32, no. 5, pp. 1–11, 2013.

[144] A. Kovnatsky, M. M. Bronstein, A. M. Bronstein, K. Glashoff, and R. Kimmel, "Coupled quasi-harmonic bases," *Computer Graphics Forum*, vol. 32, no. 2, pp. 439–448, 2013.

[145] R. Rustamov, Y. Lipman, and T. Funkhouser, "Interior distance using barycentric coordinates," *Comp Graphics Forum*, vol. 28, no. 5, pp. 1279–1288, 2009.

[146] R. Rustamov, "Interpolated eigenfunctions for volumetric shape processing," *Vis Comput*, vol. 27, no. 11, pp. 951–961, 2011.

[147] R. Litman, A. Bronstein, and M. Bronstein, "Stable volumetric features in deformable shapes," *Computers & Graphics*, vol. 36, no. 3, pp. 569–576, 2012.

[148] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on pure and applied mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[149] J. Shah, "Gray skeletons and segmentation of shapes," *Computer Vision and Image Understanding*, vol. 99, no. 1, pp. 96–109, 2005.

[150] K. Nishijima, *Fundamental Particles*. New York: W.A. Benjamin Inc., 1963.

[151] A. Grigor'yan, "Heat kernels on weighted manifolds and applications," *Cont. Math*, vol. 398, pp. 93–191, 2006.

[152] N. O'Connell, "Conditioned random walks and the rsk correspondence," *Journal of Physics A: Mathematical and General*, vol. 36, no. 12, p. 3049, 2003.

[153] J. L. Doob, *Classical potential theory and its probabilistic counterpart*. Springer-Verlag, 1984, vol. 262.

[154] S. S. Varadhan, "On the behavior of the fundamental solution of the heat equation with variable coefficients," *Communications on Pure and Applied Mathematics*, vol. 20, no. 2, pp. 431–455, 1967.

[155] M. Ovsjanikov, Q. Merigot, F. Memoli, and L. Guibas, "One point isometric matching with the heat kernel," *Comp Graphics Forum*, vol. 29, no. 5, pp. 1555–1564, 2010.

[156] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, 1994.

[157] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[158] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *The Journal of Machine Learning Research*, vol. 5, pp. 1457–1469, 2004.

[159] B. A. Olshausen *et al.*, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.

[160] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[161] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *The Journal of Machine Learning Research*, vol. 11, pp. 19–60, 2010.

[162] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and R. Jenatton, "Spams: Sparse modeling software," 2011.

[163] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[164] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.

[165] J. Bolz, I. Farmer, E. Grinspun, and P. Schröoder, "Sparse matrix solvers on the gpu: conjugate gradients and multigrid," *ACM Transactions on Graphics (TOG)*, vol. 22, no. 3, pp. 917–924, 2003.

[166] P. G. Schmitz and L. Ying, "A fast direct solver for elliptic problems on general meshes in 2d," *Journal of Computational Physics*, vol. 231, no. 4, pp. 1314–1338, 2012.

[167] J. Xia, S. Chandrasekaran, M. Gu, and X. S. Li, "Superfast multifrontal method for large structured linear systems of equations," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1382–1411, 2009.

[168] A. McAdams, E. Sifakis, and J. Teran, "A parallel multigrid poisson solver for fluids simulation on large grids," in *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Eurographics Association, 2010, pp. 65–74.

[169] N. Bell and M. Garland, "Cusp: Generic parallel algorithms for sparse matrix and graph computations," 2012, version 0.3.0. [Online]. Available: http://cusp-library.googlecode.com

[170] M.-K. Hu, "Visual pattern recognition by moment invariants," *Information Theory, IRE Transactions on*, vol. 8, no. 2, pp. 179–187, 1962.

[171] H. Jégou, F. Perronnin, M. Douze, J. Sánchez, P. Pérez, and C. Schmid, "Aggregating local image descriptors into compact codes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 9, pp. 1704–1716, 2012.

[172] M. Bartlett, "The square root transformation in analysis of variance," *Supplement to the Journal of the Royal Statistical Society*, pp. 68–78, 1936.

[173] M. F. Freeman and J. W. Tukey, "Transformations related to the angular and the square root," *The Annals of Mathematical Statistics*, pp. 607–611, 1950.

[174] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The princeton shape benchmark," in *Shape Modeling Applications, 2004. Proceedings.* IEEE, 2004, pp. 167–178.

[175] Z. Lian, A. Godil, B. Bustos, M. Daoudi, J. Hermans, S. Kawamura, Y. Kurita, G. Lavoué, H. Van Nguyen, R. Ohbuchi *et al.*, "A comparison of methods for non-rigid 3d shape retrieval," *Pattern Recognition*, vol. 46, no. 1, pp. 449–461, 2013.

[176] D. Smeets, T. Fabry, J. Hermans, D. Vandermeulen, and P. Suetens, "Isometric deformation modelling for object recognition," in *Computer Analysis of Images and Patterns*. Springer, 2009, pp. 757–765.

[177] Z. Lian, A. Godil, X. Sun, and H. Zhang, "Non-rigid 3d shape retrieval using multidimensional scaling and bag-of-features," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 3181–3184.

[178] D. Smeets, J. Keustermans, D. Vandermeulen, and P. Suetens, "meshsift: Local surface features for 3d face recognition under expression variations and partial data," *Computer Vision and Image Understanding*, vol. 117, no. 2, pp. 158–169, 2013.

[179] D. W. Thompson *et al.*, "On growth and form." *On growth and form.*, 1942.

[180] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 4, pp. 509–522, 2002.

[181] K. Siddiqi and B. B. Kimia, "A shock grammar for recognition," in *Computer Vision and Pattern Recognition, 1996. IEEE Computer Society Conference on*. IEEE, 1996, pp. 507–513.

[182] S. C. Zhu and A. L. Yuille, "Forms: a flexible object recognition and modelling system," *International Journal of Computer Vision*, vol. 20, no. 3, pp. 187–212, 1996.

[183] S.-C. Zhu, "Stochastic jump-diffusion process for computing medial axes in markov random fields," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 21, no. 11, pp. 1158–1169, 1999.

[184] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding*, vol. 89, no. 2, pp. 114–141, 2003.

[185] B. Jian and B. C. Vemuri, "A robust algorithm for point set registration using mixture of gaussians," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 2. IEEE, 2005, pp. 1246–1251.

[186] Y. Zheng and D. Doermann, "Robust point matching for nonrigid shapes by preserving local neighborhood structures," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 4, pp. 643–649, 2006.

[187] H. Guo, A. Rangarajan, S. Joshi, and L. Younes, "Non-rigid registration of shapes via diffeomorphic point matching," in *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on.* IEEE, 2004, pp. 924–927.

[188] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[189] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 2. IEEE, 2003, pp. II–264.

[190] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International journal of computer vision*, vol. 60, no. 1, pp. 63–86, 2004.

[191] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *International journal of computer vision*, vol. 65, no. 1-2, pp. 43–72, 2005.

[192] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 5, pp. 530–535, 1997.

[193] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1. IEEE, 2001, pp. 525–531.

[194] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman, "Object retrieval with large vocabularies and fast spatial matching," in *Computer Vision and Pattern Recognition, 2007. IEEE Conference on.* IEEE, 2007, pp. 1–8.

[195] T. Lindeberg, "Feature detection with automatic scale selection," *International journal of computer vision*, vol. 30, no. 2, pp. 79–116, 1998.

[196] T. Lindeberg and J. Gårding, "Shape-adapted smoothing in estimation of 3-d shape cues from affine deformations of local 2-d brightness structure," *Image and vision computing*, vol. 15, no. 6, pp. 415–434, 1997.

[197] A. Vedaldi and B. Fulkerson, "Vlfeat: An open and portable library of computer vision algorithms," in *Proceedings of the international conference on Multimedia.* ACM, 2010, pp. 1469–1472.

[198] X. Yang, X. Bai, L. J. Latecki, and Z. Tu, "Improving shape retrieval by learning graph transduction," in *Computer Vision–ECCV 2008.* Springer, 2008, pp. 788–801.

[199] X. Yang, S. Koknar-Tezel, and L. J. Latecki, "Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on.* IEEE, 2009, pp. 357–364.

[200] X. Yang, L. Prasad, and L. Latecki, "Affinity learning with diffusion on tensor product graph," 2013.

[201] X. Bai, X. Yang, L. J. Latecki, W. Liu, and Z. Tu, "Learning context-sensitive shape similarity by graph transduction," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 5, pp. 861–874, 2010.

[202] P. Kontschieder, M. Donoser, and H. Bischof, "Beyond pairwise shape similarity analysis," in *Computer Vision–ACCV 2009.* Springer, 2010, pp. 655–666.

[203] M. Donoser and H. Bischof, "Diffusion processes for retrieval revisited," pp. 1320–1327, 2013.

[204] H. Ling, X. Yang, and L. J. Latecki, "Balancing deformability and discriminability for shape matching," in *Computer Vision–ECCV 2010*. Springer, 2010, pp. 411–424.

[205] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Shock-based indexing into large shape databases," in *Computer Vision—ECCV 2002*. Springer, 2002, pp. 731–746.

[206] D. Zhang and G. Lu, "Evaluation of mpeg-7 shape descriptors against other shape descriptors," *Multimedia Systems*, vol. 9, no. 1, pp. 15–30, 2003.

[207] W.-Y. Kim and Y.-S. Kim, "A region-based shape descriptor using zernike moments," *Signal Processing: Image Communication*, vol. 16, no. 1, pp. 95–102, 2000.

[208] P. Salembier, T. Sikora, and B. Manjunath, *Introduction to MPEG-7: multimedia content description interface*. John Wiley & Sons, Inc., 2002.

[209] F. Mokhtarian and M. Bober, *Curvature scale space representation: theory, applications, and MPEG-7 standardization*. Springer Publishing Company, Incorporated, 2011.

[210] E. Baseski, A. Erdem, and S. Tari, "Dissimilarity between two skeletal trees in a context," *Pattern Recognition*, vol. 42, no. 3, pp. 370–385, 2009.

[211] X. Bai, B. Wang, C. Yao, W. Liu, and Z. Tu, "Co-transduction for shape retrieval," *Image Processing, IEEE Transactions on*, vol. 21, no. 5, pp. 2747–2757, 2012.

[212] M. A. Fischler and H. C. Wolf, "Locating perceptually salient points on planar curves," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 2, pp. 113–129, 1994.

[213] J. Brault and R. Plamondon, "Segmenting handwritten signatures at their perceptually important points," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 9, pp. 953–957, 1993.

[214] D. M. Wuescher and K. L. Boyer, "Robust contour decomposition using a constant curvature criterion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 1, pp. 41–51, 1991.

[215] C.-H. Teh and R. T. Chin, "On the detection of dominant points on digital curves," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 8, pp. 859–872, 1989.

[216] T. Pavlidis and S. L. Horowitz, "Segmentation of plane curves," *IEEE transactions on Computers*, vol. 23, no. 8, pp. 860–870, 1974.